#### (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

# (19) World Intellectual Property Organization International Bureau



# 

# (43) International Publication Date 9 August 2001 (09.08.2001)

# PCT

# (10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US];

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000) 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496.914 (CIP) Filed on 3 February 2000 (03.02.2000) US 09/560,875 (CIP) 27 April 2000 (27.04.2000) Filed on

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston,

MA 02111 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,

TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Published:

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

# NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

#### 1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

#### 2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

# 3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, \* corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10

15

20

25

30

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, butilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

10

15

20

25

30

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

#### 4. DETAILED DESCRIPTION OF THE INVENTION

# **4.1 DEFINITIONS**

5

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

5

10

25

30

35

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or 15 "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T 20 (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides. preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4<sup>20</sup> possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match  $(1 \div 4^{25})$  times the increased probability for mismatch at each nucleotide position  $(3 \times 25)$ . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

10

15

20

25

30

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

5

10

15

20

25

30

35

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134-143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO<sub>4</sub>, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

15

5

10

## 4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

35

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (*i.e.* specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

5

10

15

20

25

30

35

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

10

15

20

25

30

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

10

15

20

25

30

35

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

#### 4.3 ANTISENSE

10

15

20

25

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

10

15

20

25

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

#### 4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

. 15

20

25

30

35

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556;

Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition,

20 oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

10

# 4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

5

10

15

20

25

30

35

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

25

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl<sup>TM</sup> or Cibacrom blue 3GA Sepharose<sup>TM</sup>; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

5

# 4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

## 4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

#### 4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

5

10

15

25

30

35

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

#### 4.9 TRANSGENIC ANIMALS

10

15

· 20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

## 4.10 USES AND BIOLOGICAL ACTIVITY

10

15

20

25

30

35

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

#### 4.10.1 RESEARCH USES AND UTILITIES

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

## 4.10.2 NUTRITIONAL USES

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

# 4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

## 4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

30

35

A

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation. manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

5

10

15

20

25

30

35

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

## 4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) . as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

30

5

10

# 4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No.

15 WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

5

10

#### 4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

5

10

15

20

25

. 30

35

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus. rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by

removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed

APCs either expressing a peptide of the present invention or together with a stimulatory form of
a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the
patient. Another method of enhancing anti-viral immune responses would be to isolate infected
cells from a patient, transfect them with a nucleic acid encoding a protein of the present
invention as described herein such that the cells express all or a portion of the protein on their
surface, and reintroduce the transfected cells into the patient. The infected cells would now be
capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β<sub>2</sub> microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

15

20

25

30

35

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

5

10

15

20

25

30

35

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

## 4.10.8 ACTIVIN/INHIBIN ACTIVITY

5

10

15

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

## 4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

30

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

## 4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polypucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### 4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

ý

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen'mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

## 4.10.12 RECEPTOR/LIGAND ACTIVITY

10

15

20

25

30

35

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

30

35

5

10

15

20

25

## 4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

5

10

15

20

25

30

35

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

## 4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

#### 4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

#### **4.10.16 LEUKEMIAS**

5

10

15

20

25

30

35

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

## 4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- 10 (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

5

15

20

25

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular
   neurotoxins; and
  - (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
  - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

#### 4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

## 4.10.19 **IDENTIFICATION OF POLYMORPHISMS**

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

#### 10 4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

#### 4.11 THERAPEUTIC METHODS

15

20

25

30

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

## 4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

# 4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

5

10

15

30

35

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). 20 Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or 25 amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

## 4.12.1 ROUTES OF ADMINISTRATION

5

10

15

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

#### 4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

5

10

15

20

25

30

35

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

5

10

15

20

25

30

35

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

5

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials 5 are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass. aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above 10 mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. 15 In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, 20 hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on 25 total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other 30 agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

35

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

5

10

15

# 4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC<sub>50</sub> as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD<sub>50</sub> and ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 µg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 µg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

#### 4.12.4 PACKAGING

5

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

#### 4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain,  $F_{ab}$ , and  $F_{(ab)2}$  fragments, and an  $F_{ab}$  expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG<sub>1</sub>, IgG<sub>2</sub>, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, Proc. Nat. Acad. Sci. USA 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

# 5.13.1 Polyclonal Antibodies

5

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A. synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

# 5.13.2 Monoclonal Antibodies

5

10

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

5

10

15

20

25

30

35

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

#### 5.13.2 Humanized Antibodies

5

35

10 The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins. immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigen-15 binding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the 20 corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable 25 domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 30 2:593-596 (1992)).

# 5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al., (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse<sup>TM</sup> as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

# 5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of  $F_{ab}$  expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal  $F_{ab}$  fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an  $F_{(ab)2}$  fragment produced by pepsin digestion of an antibody molecule; (ii) an  $F_{ab}$  fragment generated by reducing the disulfide bridges of an  $F_{(ab)2}$  fragment; (iii) an  $F_{ab}$  fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv)  $F_v$  fragments.

# 5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., 1991 EMBO J., 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., <u>Science</u> 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

10

15

20

25

30

35

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

# 5.13.6 Heteroconjugate Antibodies

5

10

15

20

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HTV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

# 5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

#### 5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include <sup>212</sup>Bi, <sup>131</sup>I, <sup>131</sup>In, <sup>90</sup>Y, and <sup>186</sup>Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

#### 4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

#### 4.15 TRIPLE HELIX FORMATION

5

10

15

20

25

30.

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

#### 4.16 DIAGNOSTIC ASSAYS AND KITS

10

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that annual to a polynucleotide of the invention under such conditions, and amplifying annualed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

# 4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

#### 4.18 SCREENING ASSAYS

5

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
  - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10

15

20

25

30

35

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

#### 10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

#### 4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata et al., 1985; Dahlen et al., 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller et al., 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

5

10

15

20

25

30

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm<sub>7</sub>), is then added to a final concentration of 10 mM 1-MeIm<sub>7</sub>. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm<sub>7</sub>, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

# 4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook et al. (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*II, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviII normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme ( $CviII^{**}$ ), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a  $CviII^{**}$  digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that  $CviII^{**}$  restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

#### 4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm<sup>2</sup> and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

# 5.0 EXAMPLES

5

10

15

20

25

30

#### 5.1 EXAMPLE 1

#### Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

# 5.2 EXAMPLE 2

# 5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from <a href="http://fasta.bioch.virginia.edu">http://fasta.bioch.virginia.edu</a>) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

5

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Na	ame   SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
			976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
		*	217 225 238 271 317 404 446 469 503
İ			513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
			1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
			1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
		}	147 188 197 208 225 227-239 250 300-
			303 312 316 328-331 340 357-362 374
		,	380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550 566 571 577 585 590 594 598 634 641
	j		658 666 683 725 742 764 767 786 801
			805 810 823 826 829 831 836 841 887-
		-	923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
	1	-	1128 1142 1162 1181-1192 1199 1204
		j	1218-1219 1225 1232 1253 1267 1271-
			1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
	ł	1	566 596 663 670 746 798 816-819 876
			892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes		<del>                                     </del>	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
	·		240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
			1003 1067-1070 1118 1156 1193-1200 1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
adatt noatt	СТВСО	MIKOOI	118 129 132 138 151 158-163 182 195-
			203 215 217 238 264 269 353 384 398
			408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
		,	671-672 722 734 757-773 815 828-835
		}	874 891 898 919 926-927 976 988
			1021 1037 1041 1062 1067 1071 1080
			1083 1093 1122 1131 1185 1201 1254
			1308 1331 1335
adult kidney	GIBCO	AKD001	41 49 51 71-74 78-85 94 100-101 103-
			107 111 119-120 138 151 157 215 217-
	1	}	218 238 250 264 294 304 384 404 440
	-	1	446 454 477 504-505 509 514 518-519
		1	535 537 564 574-583 620-627 639 653
	1	1	673-675 705 753 789 831 844 851 859
		]	877 909 918 927 956 963 976 1067
achile leide are	Y	AY/7000	1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
		1	446 487 564 575 844 868 910 927 976
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514
ment innig	LOBCO	LALGOUI	0 101 111 131 10/ 402 440 490 314

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
			545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
	<b>)</b>	<b>1</b>	519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
			505 657 675 714 753 832 844 941-942
			976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101
		1	104 111 120 122-125 138 140 143-149
		İ	151 188-190 207-212 215-217 238 264
			316 384 409 440 445-446 496 504 512
			514 518-519 535 537 549-550 564 566
			571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771 815 833 842-865 868 875-876 918 926-
			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
			1124 1131 1144 1174 1224 1268 1331
			1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
addit spiech	I dibec	A51 001	294 414 446 477 504 514 534 545 549
		İ	592 722 873 883 952 976 1041-1042
			1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215
		1	238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
			592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
			147 151 164-174 213-215 238 305-307
			374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
	,		596 607 610 628-629 643-645 652 707-
	1		708 774-789 844 866-871 873 919 927
		·	952 963 976 998 1034 1042 1064 1083
			1085 1120 1132 1152 1225 1229 1268
		l l	1 1207 1210
		1	1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141
			6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346
bone marrow	Clontech	BMD004	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083
			6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592
bone marrow	Clontech	BMD004	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268
bone marrow adult colon	Clontech Invitrogen	BMD004 CLN001	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337
bone marrow	Clontech	BMD004	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332-
bone marrow adult colon	Clontech Invitrogen	BMD004 CLN001	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332- 335 365-367 392-401 440 466 470-471
bone marrow adult colon	Clontech Invitrogen	BMD004 CLN001	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332- 335 365-367 392-401 440 466 470-471 518 537 597 629 832 877 927 976 1006
bone marrow adult colon	Clontech Invitrogen	BMD004 CLN001	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332- 335 365-367 392-401 440 466 470-471 518 537 597 629 832 877 927 976 1006 1085 1117 1129-1134 1192 1202-1205
bone marrow adult colon	Clontech Invitrogen	BMD004 CLN001	6 8 37-38 52 74 77 105 111 129 132 210 317 510-511 545 549 581 598 628 638 724 766 789 844 860 868 873 919 927 952 963 968 976 1042 1111 1141 1160-1161 1229 1266 1346 111 238 282 549 1083 52 260 264 299 494 536 545 564 592 844 873 877 952 976 1042 1152 1268 1336-1337 49 51 129 132 151 205 207 238 332- 335 365-367 392-401 440 466 470-471 518 537 597 629 832 877 927 976 1006

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
			138 151 204-206 215-217 238 269 316
Į.	}		414 433 505 510 513 550 555 580 582
	1		596 675 722 745 798 814 836-841 851
	1		918 976 1041 1043 1073 1083 1131
			1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic		
of chromosome 8	Research	F60000	74 130 330
esophagus fotal basis	BioChain	ESO002	74 138 238
fetal brain fetal brain	Clontech	FBR001	441-442 927
	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208 225 271 317 319 336 359 368 405-414
}	1		1
			519 550 571 594 686 715 722 764 824 829 836 859 909 927 943 947 963 1057
	ì		1067-1068 1104 1135-1140 1162 1206-
}			1207 1235 1268 1288 1307-1308 1319
			1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
Total brain	Mvillogen	7 7 1 1 1 0 0 2	535 683 761 798 820-827 844 876 909
			963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
		,	550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
	University		69-71 74 77 79 87-90 101 107 110-111
			114 120 128-131 138 140 147 150-155
	į		197 210 215 217 225 238 312 367 384
	]		414 440 446 460 468 483 496 504-507
	[		511-515 518-519 523 533-535 537 541
	1		544-545 547-550 555-560 564 566 571
			577 582 585-586 598 636 646-647 649
			652 664 698 709-710 714 722-723 731
]	} .		735-736 746-753 761 784 798 823 829
	]		832 844 851 858-859 868 873 876 898
			927 943 949 952 963 976 984 1002
	] !		1021 1023 1040 1042 1044 1050 1083
	ĺ		1093 1116 1120 1129 1131 1144 1174
			1217 1251 1254 1256 1302 1308 1311
fotol livram and an	Columbia	ET 2002	1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
	University		111 120 129 147 207 210 215-216 238
		0	250 330 353 359 366 383-384 414 478
	ļ <b>!</b>		505 508-509 511 515-524 534-535 537 544-545 564 566 571 577 591 598 638
L <del></del>	<u> </u>		אנס אלכ ועל / וכ נוכ ססב אסב באב-האב

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			663 671 698 714 722 725 727 751 798
			851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
{	1		1131 1144 1178 1199 1233 1240-1270
<del></del>	1		1331 1340
fetal liver-spleen	Columbia	FLS003	64 535 976 1256
fetal liver	University	F7 1/001	
letai liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
		1	580 722 730 749 844 918 943 976 1051 1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
icai muscie	mvidogen	F1415001	425 535 537 577 598 614 836 857 1141
	ļ	}	1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
			225 264 316 405 422-429 488-494 496
		1	519 534-535 537 566 675 732 859 876-
		[	877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
		1	1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
			316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
			976 1001 1075 1142-1143 1171 1218
			1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
			138 145 151 188 197 207 215 238 264
•			271 294 316 367 414 440 446 466 504
			513-514 535 542-543 550 564 571 596 635 648-654 675 711-715 722-723 798
			832 872 876 883 927 976 1095 1144
	i		1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
			299 308-310 371-373 462 476 504 511-
			513 533 537 564 566 571 655-657 662
	1		683 716-720 723 752 790-803 829 832
	4		858-859 876 898 909 949 976 1045-
	1		1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
			1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University		519 566 655 714 794 918 943 976 1067
	<del></del>		1092-1093 1233
infant brain	Columbia	IBM002	311 472-473 753 1214
infont has!	University	170001	61 111 076 474 700 076 040
infant brain	Columbia	IBS001	51 111 376 474 790 876 949 1144 1204
lung , fibroblast	University	LFB001	1221
lung tumor	Strategene Invitrogen	<u> </u>	151 316 462 514 534 582 675 939 1131
imig mmor	invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156 215 217 269 280 296 337 374-375 384
			404 446 454 475-480 498 514 518-519
			522 537 545 564 577 597 653 658 705
			721-724 754-756 779 859 868 872-874
			876-877 919 927 949 951-952 959 976
		1	1002 1042 1048-1053 1076 1083 1088-
		1	1089 1131 1144-1147 1216-1218 1229
		<del></del>	1101 11 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
			634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
			147 151 212 215 218 238 252 288 312-
}			314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
		}	564 566 571 577 580 582 587-609 615
ļ			632-638 658-659 698 714 725-728 832
	1	}	836 841 859 866 873-874 882-883 918-
	i		919 927 943 952 963 976 1042 1076 1083 1090 1148 1152 1168 1195 1219-
			1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
1.000,0	Cionicon	20000	657 660 729 873 883 927 952 963 1008
	1		1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	t		919 929 939 952 976 1071 1118 1218
1424	[		1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
	[	1	217 250-256 264 297-299 305 377-378
ļ			398 446 481-486 505 512 537 545 549
Ì			571 592 725 730-733 816 829 836 844
			868 873 876-877 898 926 943 951-960
	İ	1	963 976 995 1034 1042 1048 1054-
			1055 1076 1083 1091 1093 1116-1117
induced neuron cells	Strategene	NTD001	1124 1152 1302 39 101 111 138 238 361 1225 1251
monceo nemon cens	Strategene	NIDOUI	1319
retinoid acid induced	Strategene	NTR001	74 225 976
neuronal cells	Sudiegene	, ATROOF	174 223 970
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech	PLA003	38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
	<u> </u>		1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
<del> </del>	61		1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
small intestine	Clontech	SIN001	976 1083 1120 1151 1184 8 101 147 215 259-266 446 462 505
Smatt mrestnie	Cionieca	211/001	545 592 660 789 836 866 873 927 952
}			963 967-978 1042 1120 1152 1223-
•			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
•			270 343-344 353 379 516 537 566 740
			828 927 976 979-994 1092 1153-1159
			1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
			995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
4	01	WITD 400:	963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
			750 867 874 878-881 927 963 1023
thymus	Clontech	THMc02	1083 1094-1096 1124 8 61 114 129 132 210 225 231 306
m\mas	Cioniecii	1 LITAICOS	317-319 336 340 359 380 398 446 448-
		]	463 512 519 545 554 587 598 698 724-
		}	725 789 812 836 868 873 927 947 952
	<del></del>	<del></del>	

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
			1122 1147 1177 1226-1229 1234 1311
<u></u>	4		1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
			210 217 222 253 264 271 277-286 294
	<b>!</b>		320-326 345-352 361 381-382 446 467
	İ		483 514 534 549-550 564 578 602 649
			844 882-883 927 950 956 976 1008-
			1028 1076 1083 1117-1120 1142 1163-
			1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
			545 592 611 873 883-884 927
			952 1029-1031 1042 1151-1152
			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
		1	885-886 976 1001 1032-1033
			1232

# TABLE 2

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AP036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101 -	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEO	Accession	Species	Description	Smith-	1%
D	No.	1		Waterman	Identity
NO:	1	1		Score	
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
31	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	95
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein	982	90
37	AL133215	Homo sapiens	sequence SEQ ID NO:150. bA108L7.6 (semaphorin 4G (sema domain,	687	99
			immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))		
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	AL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus annuus	hydroxyproline-rich protein	110	31
45	U82288	Caenorhabditi s elegans	Rac-like GTPase	139	70
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63	U35376	Homo sapiens	repressor transcriptional factor	485	65
64	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	785	74
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70	W75770	Homo sapiens	Human oxidoreductase YTFO3.	1144	98
71	AB011135	Homo sapiens	KIAA0563 protein	239	76
72	AB014885	Halocynthia roretzi	HrPOPK-1	813	78
73	AF045454	Cavia · porcellus	phospholipase B	955	73

ID	Accession	Species	Description	Smith-	%
	No.			Waterman	Identity
NO:	L	<u> </u>		Score	l
		musculus			
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
		norvegicus		l	
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
		<u> </u>	complex component TRAP240		
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha- l-I (hCavT3).	1357	99
79	Y14591	Human	APM-1 protein	767	100
	}	papillomaviru		1	
		s type 68		1	
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
	l	thaliana			
82	L46815	Mus	DNA binding protein Rc	895	75
		musculus			]
83	G01600	Homo sapiens	Human secreted protein, SEQ ID NO: 5681.	315	96
84	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	538	71
		<del> </del>	designated HSCOP-6.		1
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272_7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
		<u> </u>	sequence SEQ ID NO:100.		
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
		musculus	HAC3		
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94 .	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
95	Y34127	Homo sapiens		661	99
96	X03638	Rattus	Human potassium channel K+Hnov11. sodium channel protein I (aa 1-2009)	1775	92
<b>30</b>	A03036	norvegicus	sodium chamiei protein I (aa 1-2009)	1773	92
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995	99
98	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	213	38
99	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	675	48
	12 021733	norvegicus	kinase	0/3	**
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	60
	11000,000	aromo ouprono	difference at residue 58		""
102	U22829	Mus	P2Y purinoceptor	264	42
		musculus	• • • • •		
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
		<u></u>	receptor-B3.		
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
			119 SEQ ID NO:119.		<u> </u>
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia	sialic acid transporter	587	96
100	***************************************	∞li		(0.5	L
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone	182	94
	705522	1,,,	HP00631 amino acid sequence.		0.5
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	464	85
	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein	274	51.
112		Tions and	sequence SEQ ID NO:84.	201	<u> </u>
112		Homo sapiens	hexokinase 1 isoform td	301	71
112	AF016365		T1	520	75
112 113 114	AC007956	Homo sapiens	unknown	520	75
112 113 114 115	AC007956 M83738	Homo sapiens Homo sapiens	protein-tyrosine phosphatase	251	92
112 113 114	AC007956	Homo sapiens			

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:				Score	<u> </u>
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rettus	phosphatidylinositol 3-kinase	627	93
	.l	norvegicus			
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory	1646	94
			subunit precursor, PDPr		
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi	contains similarity to C2 domains	219	29
-		s elegans			
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit	693	90
126	AB021861	Mus	apoptosis signal-regulating kinase 2	153	65
		musculus	, , , , , , , , , , , , , , , , , , , ,	1	
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF208043	Mus	testis specific serine kinase-3	800	87
131	1 201/34	musculus	wous specific serine kinase-3	***	67
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73	1157	87
154	W /4602	Hollo sapicits	clone HSQEL25.	1137	67
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	5041	99
			similarity to P49205 (PID:g1345860)		
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis	DNA-damage inducible protein DDI1-like	147	55
1.10	***********	thaliana			<del> </del>
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	392	61
153	AF151859	Homo sapiens	ligand (clone 3TW). CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens		432	92
156			alternatively spliced form		. –
	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
				551	
160	AJ250425 .	Rattus norvegicus	Collybistin I	556	74

SEQ	Accession	Species	Description	Smith-	T%
ID	No.	Species	Description	Waterman	Identity
_	10.				Таспату
NO:	700000	<del> </del>		Score	<del> </del>
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
164	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
165	AF160798	Rattus norvegicus	calcium transporter CaT1	700	96
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus	NIK-related kinase	197	43
169	AF252293	musculus Homo sapiens	PAR3	596	44
170	U59429	Cricetinae	diacylglycerol kinase eta	481	82
		gen sp.			
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus musculus	semaphorin cytoplasmic domain-associated protein 3B	507	82
173	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	653	99
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus		168	55 .
		musculus	embryonic stem cell phosphatase		
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196 4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	99
			codon)		
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor	175	80
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading	985	99
107	1 200066	77	ectoenzyme	549	
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit	157	72
193	M92084	Theileria	polypeptide (MSP)GPIIb-IIIa. casein kinase II alpha subunit	364	50
		parva			
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone hj968 2.	382	49
196	AF255614	Rattus	scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
		thaliana		L	
198	AF073967	Mus musculus	olfactory receptor	316	43
		domesticus		<u> </u>	
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein	701	99
204	APOSTOCO	Uomo conier	sequence SEQ ID NO:48.	182	70
204	AF227968	Homo sapiens	SH2-B beta signaling protein		79
203	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.			Waterman Score	Identity
NO:	<del> </del>		{ovarian cancer critical region of deletion}	Store	<del> </del>
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
210	D79992	Homo sapiens	similar to Drosophila photoreceptor cell-specific	541	82
	<u> </u>	<u> </u>	protein, calphotin.		L
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C	1348	99
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	69
012	A D164046	norvegicus	neurofascin	798	56
213 214	AF154846 AF102777	Homo sapiens Mus	zinc finger protein  FYVE finger-containing phosphoinositide kinase	933	93
214	AF102/1/	musculus	F 1 VE miger-containing phospholiosidde kniese	733	} 33
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus	prostaglandin F2a receptor regulatory protein	563	78
	1	norvegicus	precursor	1	1
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus musculus	Kupffer cell receptor	567	40
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus norvegicus	mytonic dystrophy kinase-related Cdc42-binding kinase	636	96
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
		musculus	1		
226	AE000218	Escherichia coli	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus musculus	GTP-binding like protein 2	265	88
229	AF122924	Xenopus laevis	Wnt inhibitory factor-1	316	40
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754	Homo sapiens	Hurnan growth differentiation factor-12.	682	95
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific phospholipase-D.	290	100
234	W69431	Homo sapiens	Human secreted protein cw1233 3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus musculus	Embryo Brain Kinase	460	62
238	U64857	Caenorhabditi s elegans	similar to the BPTI/Kunitz family of inhibitors; most similar to tissue factor pathway inhibitor precursor (TFPI)	284	33
239	AJ250840	Mus musculus	serine/threonine protein kinase	739	63
240	AJ223472	Mus musculus	transcription elongation factor TFIIS.h	222	38
241	Y94906	Homo sapiens	Human secreted protein clone rb649_3 protein sequence SEQ ID NO:18.	353	52
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
242 243	L22022	Rattus	orphan transporter v7-3	667	93
<b>₽</b> 73	122022	norvegicus	orphian mansporon v 7-3	30,	1
244	AF016191	Rattus	potassium channel	1043	98
245	A E007266	norvegicus	conc codium coloium potessium auchan	645	98
245 246	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	497	98
246 247	Y29868 AF180475	Homo sapiens	Human secreted protein clone pp325_9.  Not4-Np	188	83
44 I		Homo sapiens Homo sapiens	Human secreted protein (clone yal-1).	690	99
248	Y17227				

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	ļ	<del></del>		Score	<u> </u>
	1710005	sexta	protein SCLP	1.2.	ļ
250	AF192756	Kaposi's	Orf73	134	34
	ł	sarcoma-		1	ì
		associated		i	l
	-	herpesvirus		ļ	<u> </u>
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (ethb0018f2 product).	469	100
253	L46815	Mus	DNA binding protein Rc	251	67
		musculus	<u>                                     </u>		<u> </u>
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066	Mus	Citron-K kinase	1201	98
		musculus		l	<u> </u>
256	G02491	Homo sapiens	Human secreted protein, SEQ ID NO: 6572.	460	100
257	Z12841	Oryctolagus	Phospholipase	368	80
		cuniculus		j	
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus	L-periaxin	430	72
	1	musculus	•		1
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
		norvegicus		1	
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
	- 11002	120	(GPCR).	1 334	
266	U27269	Mus	sodium glucose cotransporter	204	56
		musculus	Contract descriptions	{ -4.	
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus	putative taste receptor TR1	209	39
-00	14 12/303	norvegicus	padaro dallo rocopio. Trei	200	
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fc-gamma receptor	129	26
2,0	77,0402	pyogenes	1 o-gamma recopior	127	120
271	AB009883	Nicotiana	KED	109	26
-/.	1.00000	tabacum	Temporal Control of the Control of t	100	120
272	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
	1.4.157507	musculus	VI BIV commin receptor protein Bercee	055	1"
273	L34938	Rattus	ionotropic glutamate receptor	460	86
275	234730	norvegicus	i ionoropio grammate receptor	1 700	] 00
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent	188	74
	1.0000.24	Tromio sapions	Expressed Protein LIKE PUTATIVE protein)	100	' '
		i	(isoform 1)	ļ	ł
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	173	94
213	AL 203333	Tionio sapiciis	APOLLON	1 1/3	177
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis	Contains PF 00069 Eukaryotic protein kinase	157	43
219	AC008073	thaliana	domain.	137	43
200	1/02720			101	72
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	497	84
283	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
204	1,000	musculus	<u></u>	-	1.00
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate	647	100
			reading frame protein.		ļ
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein	300	90
			sequence SEQ ID NO:26.		
286	AF016411	Homo sapiens	KCNA3.1B	137	100
287	W89253	Homo sapiens	Human ALP.	688	97
288	AF112886	Bos taurus	differentiation enhancing factor 1	750	96
	AF113131	Homo sapiens	host cell factor homolog LCP	367	44
289					1 4 4 4
289 290 291	U52111 AF026504	Homo sapiens	plexin-related protein	698	100 89

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1		Waterman	Identity
NO:	<del> </del>	+		Score	ļ
292	AF102854	norvegicus Rattus		124	ļ
292	AF 102854	norvegicus	membrane-associated guanylate kinase-	124	53
293	X99211	Drosophila	interacting protein 2 Maguin-2	<del>                                     </del>	<del> </del>
293	X99211	melanogaster	ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	185	94
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577_1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	182	97
299	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	605	69
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyl-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases; Method: conceptual translation supplied by author	428	72
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus musculus	membrane glycoprotein	199	41
308	AF255614	Rattus norvegicus	scaffolding protein SLIPR	639	88
309	579463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium discoideum	calcium binding protein	151	36
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	744	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789 .	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918)	197	38
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and GENEWISE)	278	38
316	U70209	Mus musculus	polycystic kidney disease 1 protein	165	38
317	AF109643	Rattus norvegicus	coxsackie-adenovirus-receptor homolog	223	38
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma vivax	activated protein kinase C receptor homolog	141	38
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate receptor	232	97
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	306	88
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37)	214	97
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable region	581	80
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP- 107 SEQ ID NO:107.	1127	100
333	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	320	98
334	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	327	93

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
		<del>                                     </del>	similarity to P49205 (PID:g1345860)		<del> </del>
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	1111	67
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia coli	transmembrane protein	232	100
345	M60177	Escherichia coli	enterobactin	759	99
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus musculus	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+- activated potassium channel	463	80
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	577	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVH1	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus norvegicus	phospholipase C delta-4	649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368 369	AL021366 U70932	Homo sapiens Peromyscus	clCK0721Q.3 (Kinesin related protein) reverse transcriptase	3387 92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172	55

SEQ	Accession	Species	Description	Smith-	%
ID	No.	,		Waterman	Identity
NO:				Score	
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus musculus	GTP binding protein	1456	91
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium discoideum	protein tyrosine kinase	115	44
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228 6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates syndactylus	dopamine receptor D4	105	35
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1446	96
			apoint mid5ott + (01 AO+)))	1	1

SEO	Accession	Species	Description	Smith-	1%
D O	No.	- Production		Waterman	Identity
NO:				Score	1000000
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO 1	6268	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ	1961	99
	1	Tromo suprons	ID NO. 191.	1701	"
426	AB009288	Homo sapiens	N-copine	635	98
427	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
430	104441	Homo sapiens	associated protein.	2074	100
431	G02850	17ana aasiana	Human secreted protein, SEQ ID NO: 6931.	702	105
432	G04067	Homo sapiens		723	95
433		Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon	extensin-like protein	613	48
424	17740261	esculentum			<u> </u>
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52
438	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	168	56
439	X14766	Homo sapiens	GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rathus	transmembrane receptor UNC5H1	817	93
		norvegicus		***	1 -2
450	AF081249	Homo sapiens	JAW1-related protein MRVI1A long isoform	4568	99
451	AC005498	Homo sapiens	R31665 1	316	62
452	M60235	Homo sapiens	granule membrane protein-140	464	73
453	AB036706	Homo sapiens	intelectin	730	88
454	G00918	Homo sapiens	Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
	12203.	Tromo suprons	(CIRP-1).	172	67
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
150	130.03	Tiomo sapions	gene 62.	100	40.
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
150	1117133	falciparum	5-andgen precursor	110	30
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	1100 (00	<del></del>			
700	¥02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	149	43
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
701	117702	Tromo subjetts	gene 17.	104	34
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein	135	47
702	1 23003	raomo sabiens		122	1 47
463	X84960	Triticum	sequence SEQ ID NO:16.	100	22
COP	A04500		low molecular weight glutenin	109	33
464	W19919	aestivum	Viscon V - 1 (Islands	1001	0.5
465	AF189764	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
403	AF 169/04	Mus	alpha/beta hydrolase-1	502	59
ACC	1102500	musculus	-40	161	
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	99
460	000000	<del>                                     </del>	gene 77.		
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus	neurotoxin homologue	118	47
		musculus		·	
471	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75
472	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	72	57

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	- Potos	Downpuon	Waterman	Identity
NO:	1	1		Score	Identity
-	<del>                                     </del>		gene 62.	1 50010	<del> </del>
473	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	328	100
474	Y07007	Homo sapiens	Breast cancer associated antigen precursor	1013	97
1			sequence.	1 .0.5	1"
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	202	60
			clone HTDAD22	1	1
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	3427	92
		musculus			1
480	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
481.	W87701	Homo sapiens	A human membrane fusion protein designated	221	77
		<u> </u>	SYTAX1.		
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	131	39
483	AF210651	Homo sapiens	NAG18	124	59
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	50
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	70
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	149	73
<u> </u>		ļ	3		
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	. J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	197	67
	102055	Tromo supions	clone HTDAD22.	197	67
495	AC005175	Homo sapiens	R31449 3	889	94
496	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
497	AB030237	Canis	D4 dopamine receptor	90	48
		familiaris			1
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	65
499	U70935	Peromyscus	reverse transcriptase	213	52
		maniculatus	•		
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124_3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dh1135_9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	115	33
500	47.160156	ļ.,	NO:180.		
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light	184	92
509	1142260	7	polypeptide kinase))		
309	U43360	Peromyscus maniculatus	reverse transcriptase	97	62
510	G03789		II OTO ID NO COCO		
511	W79092	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	117	63
512 .	AF010144		Human secreted protein dn740 3.	1058	100
513	AJ133439	Homo sapiens Homo sapiens	neuronal thread protein AD7c-NTP GRIP1 protein	205	64
514	AE003456	Drosophila	CG6393 gene product	2151	100
214	110007430	melanogaster	CO0333 Belie broduct	259	42 .
515	Z17206	Xenopus	p46XlEg22	128	40
		laevis	PANYORES	120	40
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	92	40
518 ·	AF151083	Homo sapiens	HSPC249	444	98
519	S80864	Homo sapiens	cytochrome c-like polypeptide	318	
520	X92485	Plasmodium	pval	170	61
		vivax	k	1/0	01
	<u> </u>			I	

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Description	Waterman	Identity
NO:	140.	1		Score	localuty
521	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	159	59
522	AF121857				
		Homo sapiens	sorting nexin 7	259	40
523	G02654	Homo sapiens	Human secreted protein, SEQ ID NO: 6735.	82	37
524	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	253	73
525	AF119851	Homo sapiens	PRO1722	162	57
526	Y27761	Homo sapiens	Human secreted protein encoded by gene No. 47.	154	57
527	G02707	Homo sapiens	Human secreted protein, SEQ ID NO: 6788.	70	45
528	U47924	Homo sapiens	C8	1112	86
529	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	84	45
530	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	111	60
531	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	92	65
532	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	75	29
533	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	182	48
534	AF068286	Homo sapiens	HDCMD38P	861	100
535	U07707	Homo sapiens	epidermal growth factor receptor substrate	228	60
536	G01955	Homo sapiens	Human secreted protein, SEQ ID NO: 6036.	484	75
537	AF219232	Gallus gallus	gin-induced kinase	206	53
538	AF135022	Homo sapiens	mediator	128	100
539	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	141	59
540	AF016430	Caenorhabditi s elegans	contains similarity to a BR-C/TTK domain	853	39
541	AC003093	Homo sapiens	OXYSTEROL-BINDING PROTEIN; 45% similarity to P22059 (PID:g129308)	408	66
542	M29487	Homo sapiens	integrin alpha subunit precursor	517	81
543	AF102530	Mus musculus	olfactory receptor F3	327	73
544	Y73431	Homo sapiens	Human secreted protein clone yb186_1 protein sequence SEQ ID NO:84.	386	100
545	AE004833	Pseudomonas	probable TonB-dependent receptor	279	42
£46	002702	acruginosa	77		ļ
546 547	G03793 Y69192	Homo sapiens	Human secreted protein, SEQ ID NO: 7874.	264	53
		Homo sapiens	A human monocyte-macrophage apolipoprotein B receptor protein.	1772	67
548	Y91493	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:166.	176	100
549	G01571	Homo sapiens	Human secreted protein, SEQ ID NO: 5652.	777	99
550	AF044588	Homo sapiens	protein regulating cytokinesis 1; PRC1	1953	88
551	Y29332	Homo sapiens	Human secreted protein clone pe584_2 protein sequence.	1224	94
552	X98330	Homo sapiens	ryanodine receptor 2	24621	99
553	Y42782	Homo sapiens	Human UC Band #331 protein.	684	95
554	AB025258	Mus musculus	granuphilin-a	501	41
555	AJ010346	Homo sapiens	RING-H2	1468	100
556	W92388	Homo sapiens	Human TR-interacting protein S239a.	538	92
557	AF119851	Homo sapiens	PRO1722	175	59
558	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	183	32
559	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	319	68
560	D86214	Mus musculus	Ca2+ dependent activator protein for secretion	1010	93
561	AF187325	Canis familiaris	melanoma antigen	287	55
562	AJ001981	Homo sapiens	OXAIL	2512	99
563	Z17238	Rattus norvegicus	glutamate receptor subtype delta-1	338	66
564	W30638	Homo sapiens	Partial human 7-transmembrane receptor HAPO167 protein.	371	100
565	AC005620	Homo sapiens	R33590_1	467	97
566	Y99358	Homo sapiens	Human PRO1772 (UNQ834) amino acid sequence SEQ ID NO:63.	1138	78
567	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	1002	58
568	AF151043	Homo sapiens	HSPC209	798	100
	LT 121043	Troung Sahienz	1101 0207	170	100

SEQ	Accession	Species	Description	Smith-	%
ID `	No.		•	Waterman	Identity
NO:				Score	
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100
	10,050	lious suprem	sequence.		1
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574			seven transmembrane domain orphan receptor	1883	99
	AB037108	Homo sapiens		836	100
575	D43949	Homo sapiens	This gene is novel.		1
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
579	AK025116	Homo sapiens	unnamed protein product	201	70
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ	77	70
	<u> </u>	<u> </u>	ID NO:388.		1
581	AF196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
	<u> </u>	familiaris			
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	268	85
			Antigen)	l	
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
	1.2.2.2.	musculus	p	1	1
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	329	81
	11.00027	TIOMO Supioid	HPMBQ32.	"-"	1 0.
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted	110	43
	130709	Homo sapicus	protein.	110	1 43
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
392	1330/3	Homo sapiens	_	1309	1 32
593	Y53051	Homo sapiens	polypeptide.  Human secreted protein clone dd119_4 protein	1112	97
293	123021	rionio sapiens		1112	31
701	1405660		sequence SEQ ID NO:108.	763	79
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.		
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus	COP1 protein	2215	95
		musculus			
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus	putative secreted protein ZSIG37	143	40
	<u> </u>	musculus			
599	AF119855	Homo sapiens	PRO1847	236	76
600	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2	1333	93
			Antigen)		
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
				1377	99
608	W69341	Homo sapiens	Secreted protein of clone CG279_1.		
609	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone	339	82
<del></del>	100000	<del>  ,,</del>	НРМВQ32.	ļ.,,,	162
610	¥27868	Homo sapiens	Human secreted protein encoded by gene No.	116	62
	L	<u> </u>	107.		1.00
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	195	59
		1	clone HTDAD22.		<u></u>
614	M87053	Rattus	lens membrane protein	450	84
	1	norvegicus			
615	AC004232	Homo sapiens	FPM315	163	37
			Human secreted protein, SEQ ID NO: 6065.		79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Y01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650 651	G02872 Y32199	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6953.  Human receptor molecule (REC) encoded by	173 1012	78 100
652	AB032909	Hylobates	Incyte clone 2022379. dopamine receptor D4	122	32
752	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	agilis		<u> </u>	
653	AK021848	Homo sapiens	unnamed protein product	186	69
654	W73411	Homo sapiens	Human secreted protein encoded by Gene No. 15.	57	37
655	L22455	Rattus norvegicus	mu opioid receptor	116	34
	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
			Human appeared messain CEO ID NO. 6426	459	97
657	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.		
656 657 658	G02345 W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
657	G02345		Secreted protein encoded by gene 94 clone		

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:				Score	
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	375	43
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein (hs6M1-21))	480	55
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens Mus	Human cerebral protein-1.	182	87
	AB030184	musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KIAA0240) .	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88 .
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	) Spacial	Description	Waterman	Identity
NO:	1	1 :	·	Score	lading
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
			ligand (clone 2DD).	1	1
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	59
		es cerevisiae			
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
		norvegicus		<u> </u>	
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota	olfactory receptor	615	83
714	00266	marmota			<u> </u>
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
716	AB033062 G00577	Homo sapiens	KIAA1236 protein	1380	100
717	Y96864	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4658. SEQ. ID. 37 from WO0034474.	835	73 99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
117	047334	Tionio sapicais	receptor beta4 subunit	3/8	"
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
	1		designated HSCOP-6.		'
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter,	111	41
	ļ	tigrinum	NBC		
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
		musculus	protein 3A	l	
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus	potassium channel	370	100
		norvegicus			
727	AB029559	Rattus .	BATI	139	35
700	3700503	norvegicus	LYONY Y		1
728 729	Y28503 AJ011415	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
730	Z93096	Homo sapiens Homo sapiens	plexin-B1/SEP receptor bK390B3.1 (manic fringe (Drosophila)	729 142	56
/30	293090	Homo sapiens	homolog)	142	68
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor	675	99
,_,	3,000	Troute supre.s	homologue Vanilrep1.	0,5	· .
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	99
734	AE000493	Escherichia	putative transport protein	592	97
		coli			
735	AL033379	Homo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	2173	99
		}	(rhodopsin family) protein similar to high-		1
	<del>                                     </del>		affinity lysophosphatidic acid receptor homolog)	L	
736	AF132599	Homo sapiens	RANTES factor of late activated T lymphocytes-	245	56
737	V55010	Yama		002	100
738	X55019 X91906	Homo sapiens	acetylcholine receptor delta subunit	883	99.
739	AB026116	Homo sapiens Homo sapiens	voltage-gated chloride ion channel organic anion transporter 4	1978 1444	100
740	D00570	Mus	open reading frame (196 AA)	83	98
770	1 500370	musculus	open reading traine (150 AA)	ا م	24
741	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32	448	95
	1		clone HLTCJ63.		
746	W57260	Homo sapiens	Human semaphorin Y.	2414 .	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA	968	65
			from plasmid pGCS2232.		
748	Y94935	Homo sapiens	Human secreted protein clone yd218_1 protein	622	100
			sequence SEQ ID NO:76.		
749	AL022238	Homo sapiens	dJ1042K10.5 (novel protein)	314	85
750	G03889	Homo sapiens	Human secreted protein, SEQ ID NO: 7970.	391	87

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
751	AB025258	Mus musculus	granuphilin-a	773	41
360	V/5220/		IV.	900	99
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.		
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus norvegicus	vasopressin receptor	979	68
756	AF190501	Homo sapiens	leucine-rich repeat-containing G protein-coupled receptor 6	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43 clone HTADX17.	461	87
758	Z22535	Homo sapiens	ALK-3	439	98
759	R04932	Homo sapiens	Interferon-gamma receptor segment from clone 39 responsible for binding the target.	564	97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus musculus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus musculus	interleukin-1 delta	309	45
768	AF121975	Mus musculus	odorant receptor S18	268	62
760	AB008515		RanBPM	611	57
769 770	Y09945	Homo sapiens Rattus norvegicus	putative integral membrane transport protein	458	50
<del></del>	A 700/201		AD026	688	99
771 772	AF226731 Y27132	Homo sapiens Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
				213	45
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.		
779 780	AL357374 AF100346	Homo sapiens Homo sapiens	bA353C18.2 (novel protein) neuronal voltage gated calcium channel gamma- 3 subunit	232 1434	100 89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	30
784	W49042	Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOV1	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
	<del></del>				
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP95	1037	100
794	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	124	62
795	Y69384	Homo sapiens	Amino acid sequence of a 14274 receptor protein.	119	100
796	W40215	Homo sapiens	Human macrophage antigen.	1358	99

SEO	Accession	Species	Description	Smith-	1%
ID	No.	open	2000 paul	Waterman	Identity
NO:				Score	
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
798	AF159615	Homo sapiens	FGF receptor activating protein 1	461	98
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	797	99
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CRI protein.	11963	97
1		(human)	•		
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter LAT2	1364	90
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
L	1	<u> </u>	clone HOVBA03.	l	
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	784	100
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	358	100
	i		gn114_1.	ì	
817	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
818	AF151800	Homo sapiens	CGI-41 protein	1106	95
819	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
820	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored protein GPI-122.	4897	99
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma- 2 subunit	1105	100
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded from gene 28.	1540	100
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by gene 24 SEQ ID NO:147.	541	98
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262 .	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828_1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi s elegans	glycine-rich	85	36
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	998	75
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543_3.	1105	67
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
842	AF036717	Homo sapiens	FGFR signalling adaptor SNT-1	2629	99
843	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein sequence SEQ ID NO:114.	1089	100
844	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	357	69
845	AF151810	Homo sapiens	CGI-52 protein	1443	88
846	X83378	Homo sapiens	putative chloride channel	1620	99
847	AC004883	Homo sapiens	similar to general transcription factor 2I; similar	655	96

SEQ	Accession	Species	Description	Smith-	<b>1%</b>
$\mathbf{D}_{i}$	No.		•	Waterman	Identity
NO:				Score	
		<del>                                     </del>	to AF038969 (PID:g2827207)		<del>†</del>
848	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
0.,	110005507	Tromo suprems	P34984 (PID:g464305)	100	1.0
850	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
851	AF124490				
		Homo sapiens	ARF GTPase-activating protein GIT1	3415	98
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1245	99
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
			Human PKO1083 protein sequence.		
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	¥41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
862 .	Y25776	Homo sapiens	Human secreted protein encoded from gene 66.	895	99
863	Y74188	Homo sapiens	Human prostate tumor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532				1
		Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from gene 43.	388	88
869	J00123	Homo sapiens	preproenkephalin (	1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1048	98
_			gene 25 SEQ ID NO:305.		
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 ·	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
873	AF161382	Homo sapiens	HSPC264	1124	99
874	G03412	Homo sapiens	Human secreted protein, SEQ ID NO: 7493.	464	100
875	Y27572	Homo sapiens	Human secreted protein encoded by gene No. 6.	573	96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus	neurexophilin	1448	98
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted protein.	321	100
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865				
		Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
885	AF070661	Homo sapiens	HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887	X92744	Homo sapiens	hBD-1	375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone cn621 8.	994	94
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595 ·	99
890					
	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein sequence SEQ ID NO:34.	537	100
895	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
896	M24110	<del></del>		481	
897		Homo sapiens	GOS19-2 peptide precursor		100
	Z68747	Homo sapiens	imogen 38	2018	99
898 899	AF186112 AF225420	Homo sapiens	neurokinin B-like protein ZNEUROK1	619	100
	. ALTIGATA	Homo sapiens	AD025	734	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162_1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929 930	AF110399 AF210317	Homo sapiens Homo sapiens	elongation factor Ts facilitative glucose transporter family member	1666 2763	100 99
	1,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	· ·	GLUT9		100
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931 274	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.		100
933	U47924	Homo sapiens	B-cell receptor associated protein	146 <del>9</del> 529	100 93
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	196	63
935 936	AB039371 X56385	Homo sapiens	mitochondrial ABC transporter 3	1064	100
930	X30363	Canis familiaris	rab8	1004	100
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	op		Waterman	Identity
NO:				Score	
951	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1314	100
952	Y36111	Homo sapiens	Extended human secreted protein sequence, SEQ	402	70
		•	ID NO. 496.		
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRI	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202_3 protein	587	100
	100002	I I I I I I I I I I I I I I I I I I I	sequence SEQ ID NO:110.	1	1
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964	AF078859	Homo sapiens	PTD004	2089	100
965	AB020315	Homo sapiens	homologue of mouse dkk-1 gene: Acc#	1466	100
			AF030433		
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	993	99
968	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429 `	Homo sapiens	hyperpolarization-activated cyclic nucleotide	6295	100
			gated cation channel hHCN4		1
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein	3348	99
		•	complex component TRAP80		1
979	AF044201	Rattus	neural membrane protein 35; NMP35	1570	92
		norvegicus	·		
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein 1	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-	1553	99
•			encoded protein.		
983	Z56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	172	70
			gene 17.		
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca	juvenile hormone esterase binding protein	226	32
		sexta			
989	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
990	AF204159	Homo sapiens	potassium large conductance calcium-activated	1486	100
			channel beta 3a subunit		<u> </u>
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi	VM106R.1	327	40
	L	s elegans			<u> </u>
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995		Homo sapiens	corin	5811	99
996	AF133845			1 4000	100
,,,	AF133845 AF117756	Homo sapiens	thyroid hormone receptor-associated protein	4999	
	AF117756		complex component TRAP150		
997	AF117756 W62066	Homo sapiens	complex component TRAP150  Human stem cell antigen 2.	284	93
	AF117756		complex component TRAP150  Human stem cell antigen 2.  Human secreted protein sequence SEQ ID		
997 998	AF117756 W62066 Y87173	Homo sapiens Homo sapiens	complex component TRAP150  Human stem cell antigen 2.  Human secreted protein sequence SEQ ID NO:212.	284 725	93
997 998 999	W62066 Y87173 Y13379	Homo sapiens Homo sapiens Homo sapiens	complex component TRAP150  Human stem cell antigen 2.  Human secreted protein sequence SEQ ID NO:212.  Amino acid sequence of protein PRO263.	284 725	93 100 99
997 998	AF117756 W62066 Y87173	Homo sapiens Homo sapiens	complex component TRAP150  Human stem cell antigen 2.  Human secreted protein sequence SEQ ID NO:212.	284 725	93

SEQ	Accession	Species	Description	Smith-	1%
D T	No.	- France		Waterman	Identity
NO:				Score	lucauty
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No.	2150	100
1 .005	1175420	Tromo sapions	24.	2130	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1005	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ	824	99
i		1	ID NO. 382.	Ĺ	39
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by	768	100
1023	AE000660	Homo sapiens	gene 93 SEQ ID NO:362. hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell	830	
		<u> </u>	line.		100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMPI	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040 -	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like	1704	50
			growth factor binding protein, acid labile subunit))		
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81	176	100
			clone HMWDN32.		
1048 1049 ·	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
1050	AF097518	Homo sapiens	liver-specific transporter	2820	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazzpine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064 ·	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
1067	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	85
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077 1078	L25899 Y91447	Homo sapiens Homo sapiens	ribosomal protein L10	332	76
			Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079 1080	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527 L13802	Homo sapiens	Na/PO4 cotransporter homolog ribosmal protein small subunit	269	100
1083	W75098	Homo sapiens Homo sapiens	Human secreted protein encoded by gene 42	499	80
			clone HSXBI25.	143	81
1084 1085	G03564 G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 7645. Human secreted protein, SEQ ID NO: 8144.	83	51
1086	AF090942	Homo sapiens	PRO0657	88 124	43
1087	G00517	Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	124	41
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 4398.  Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	114	32
1091	S72304	Mus sp.	LMW G-protein	146 .	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123 5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	59
			clone HTDAD22.		L

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	] *		Waterman	Identity
NO:				Score	
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus	ribosomal protein L28	128	69
L		musculus			1
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presenilins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
			ID NO. 155.		
1121	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	122	65
1122	Y00337	Homo sapiens	Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by gene 49 SEO ID NO:170.	542	100
1134	AB017908 ·	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens		312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12	623	100
1142	G04091	Homo sapiens	(PPRG-12).  Human secreted protein, SEQ ID NO: 8172.		
1142	AB030235	Canis	D4 dopamine receptor	113	38
1144		familiaris		89	48
	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEQ ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	110.			Score	испилу
		s elegans	cerevisiae zinc resistance protein		
1150	G03438	Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1151	G01003	Homo sapiens	Human secreted protein, SEQ ID NO: 5084.	181,	80
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	198	63
1153	X88799	Oryza sativa	DNA binding protein	95	41
1154	D85245	Homo sapiens	TR3beta	155	96
1155	R74272	Homo sapiens	Tumour suppressor protein, p53.	341	87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	99	41
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP-	113	31
11.25	13164500		29 SEQ ID NO:29.		
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286 Y14482	Homo sapiens	HC6	134	64
		Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	· Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocysteine associated protein	249	62
		norvegicus	, , ,		1
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	107	71 .
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7539.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein	187	70
1192			complex component TRAP230		
	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens CI542_2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID NO:175.	1089	89
1200	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	154	57

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	No.	1		Score	locality
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapiens	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209 1210	AF099137 AF205718	Homo sapiens Homo sapiens	MaxiK channel beta 2 subunit hepatocellular carcinoma-related putative tumor	475	95 79
1210	Y27868		suppressor	224	79
1211	12/808	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	10
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-2) polypeptide.	725	100
1221	W96745	Homo sapiens	High affinity immunoglobulin E receptor-like protein (IGERB).	650	98
1222	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 160.	135	31
1223	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733 AF099973	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
		Mus musculus	schlafen2	333	56
1228 1229	G01218 AF217188	Homo sapiens Mus	Human secreted protein, SEQ ID NO: 5299.	155 801	81
		musculus	YPIB	·	63
1230	AF176813	Homo sapiens	soluble adenylyl cyclase	275	100
1231	X98333	Homo sapiens	organic cation transporter	1704	100
1232	W74955	Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24.	212	53
1233	Y94940	Homo sapiens	Human secreted protein clone yi62_1 protein sequence SEQ ID NO:86.	526	100
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1236	G01459	Homo sapiens	Human secreted protein, SEQ ID NO: 5540.	417	100
1237	AF000018	Homo sapiens	adapter protein	164	84
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27_1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	<u></u>		<u>}</u>	Score	
			ID NO. 160.		
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
		norvegicus			ł
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
	i .	•	gene 27 SEO ID NO:131.	1	1
1251	L15313	Caenorhabditi	putative	858	59
	]	s elegans	) Principal Control of the Control o	1 000	) "
1252	Y29338	Homo sapiens	Human secreted protein clone it217 2 alternate	278	75
	12,550	Troine supiens	reading frame protein.	278	1'3
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	
1255	G01818	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.  Human secreted protein, SEQ ID NO: 5899.		83
1256	AF286368			253	91
1257		Homo sapiens	eppin-1	222	54
	AF220264	Homo sapiens	MOST-1	87	93
1258	G02227	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1259	Y07970	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
	<u> </u>		from gene 26.		]
1260	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain	986	100
	<del></del>		ligand (clone 3TW).	<u> </u>	1
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	172	36
1262	U28369	Homo sapiens	semaphorin V	237	67
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	288	71
	1		sequence.	1	
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
	/		ID NO:2.	1 -23	"
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus	phosphatidylinositol 5-phosphate 4-kinase	859	95
	12 030330	norvegicus	gamma	0.59	33
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2		
1209	AI 190004	musculus	LMBRZ	552	76
1270	AL050332		distance in the second	000	-
12/0	AL030332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
1271	G02126	77	GTPase-activating protein p135 SynGAP)		<u> </u>
		Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform .	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme	1280	100
	l .		similar to acetyl-coenzyme A synthethase		ļ
			(acetate-coA ligase))		
1274	AF064748	Mus	S3-12	3523	61
		musculus			
1275	D17554	Homo sapiens	TAXREB107	377	78
1276	Y30715	Homo sapiens	Amino acid sequence of a human secreted	643	90
		L	protein.		1
1277	AF146760	Homo sapiens	septin 2-like cell division control protein	707	100
1278	Y05069	Homo sapiens	Human PIGR-2 protein sequence.	281	46
1279	X59668	Oryctolagus	aorta CNG channel (rACNG)	267	85
-		cuniculus		-0,	ا "
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	489	98
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 7492.	120	43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1		100
1283	AF117814	Mus	odd-skipped related 1 protein	1635	
203	A 11/014	1	oda-sythhed teistied t bioteiti	357	98
1284	U87318	musculus	N-D0 2		-
1204	08/318	Xenopus	NaDC-2	535	60
205	ADOCTO	laevis			
1285	AF061346	Mus	Edp1 protein	452	68
		musculus			
1286	AB030182	Mus	contains transmembrane (TM) region	582	68
	<u> </u>	musculus			
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
		construct			
1288	AF254411	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
1289	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98

SEQ	Accession	Species	Description	Smith-	1%
D D	No.	0,11.00	1 - 555, Q.202	Waterman	Identity
NO:				Score	100
1290	AF038563	Homo sapiens	membrane associated guanylate kinase 2	523	100
1291	AF034837	Homo sapiens	double-stranded RNA specific adenosine	468	100
	1.2.00	110.000 000,000	deaminase	100	100
1292	M15888	Bos taurus	endozepine-related protein precursor	937	87
1293	AB010692	Arabidopsis	ATP-dependent RNA helicase-like protein	636	45
12/3	1,0010072	thaliana	i ATT -dependent KNA tienease-like protein	030	43
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
1295	W67828	Homo sapiens	Human secreted protein encoded by gene 22		
1273	W07020	riomo sapiens	clone HFEAF41.	504	98
1296	AC004832	- IIama amiana		1640	1
		Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus	cysteine rich hair keratin associated protein	575	70
		cuniculus			
1298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89 13 amino acid	3916	99
	1 .0,515	1 TOTAL SAPICAS	sequence.	3510	1 "
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens			
1304	AF148509	Troing sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
	1	Homo sapiens	alpha 1,2-mannosidase	602	98
1306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying	332	98
1308	AF033120	Homo sapiens	protein SEQ ID NO:1.	1 240	<del> </del>
1309			p53 regulated PA26-T2 nuclear protein	348	52
	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus	arsenite inducible RNA associated protein	688	70
1110	3753340	musculus			
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus	Elf-I	806	92
1310	015017	musculus	Di1-1	800	92
1319	U82598	Escherichia	ferric enterobactin transport protein	768	100
	ļ	coli			
1320	D90892	Escherichia	SORBITOL-6-PHOSPHATE 2-	709	100
	į	coli	DEHYDROGENASE (EC 1.1.1.140)		
			(GLUCITOL-6- PHOSPHATE		
	Į		DEHYDROGENASE) (KETOSEPHOSPHATE		Į.
	<u> </u>		REDUCTASE).	_	1
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41	601	92
			clone HPBCJ74.	<u></u>	
1322	AJ276101	Homo sapiens	GPRC5B protein	466	93
1323	AJ276101	Homo sapiens	GPRC5B protein	504	97
324	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1584	100
325	U91561	Rattus	pyridoxine 5'-phosphate oxidase	1277	89
		norvegicus	•		
326	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
327	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by	1531	90
		, , ,	Incyte clone 2825826.	]	1
328	AF151048	Homo sapiens	HSPC214	657	85
329	Y10530	Homo sapiens	olfactory receptor	1645	100
330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
331	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform	3591	99
	1	Tromo sapiens	NaPi-3b	3371	77
				<b></b>	
	Y13583	Homo sapiens	G-protein coupled receptor	2171	100

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339 ·	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

## TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence 337	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC
		^		337	<b>.</b>	HWPQAPHRA***GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMRQMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH\RTPI*R SCNYTLALILFL
4	1354	A	74		292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	<b>A</b>	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFITAGVTMGFF MTGYLPLGFEFAVELSYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	Α.	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLI/RE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	(	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	}	1	ł	peptide		/-possible nucleotide deletion, \-possible
L				sequence	<u></u>	nucleotide insertion
ì		1				WSHEGEILQAFRGHQGRGIRALAAHERQAWV
ļ	]	ļ			]	ITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQ
ì		İ				VP**ARYTQGCDSGWLLATAGSD*YRGPVSL
1		1	1		ł	*RRGQVLGAAARG*TFPVLLPAGGSSWSRGL
ĺ						RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS
<u> </u>						WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF
ļ.	1		ì			LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL
}		ì	1		İ	LI/CRRL*SDSCMRAALNDRGLYQVLILDGLV
						QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN
10	1262	<u> </u>	1.00		1010	KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY
	1	]	1			FEEVQRLRFEVHDISSNHNGLKEADFLGGME
	l		[			CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA
						EELSGNDDYVELAFNARKLDDKDFFSKSDPF
1	l	ł	1			LEIFRMNDDATQQLVHRTEVVMNNLSPAWK
İ			<b>,</b>			SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK
ļ						HDFIGEFTSTFKEMRGAMEGKQVQWECINPK
j						YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI
ļ	ŀ					MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP
		ŀ				YQPNEYLKALVAVGEICQDYDSDKMFPAFGF
<b>{</b>	l	ł	}			GARIPPEYTDSHDFAINFNEDNPECAGIQGVV
1	<u> </u>	l				EAYQSCF\PKAPTFTGPTNICPHSSRKVAKFRR
1	1	)				SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP
<u></u>	ļ. <u></u> .					DNPGGHFV
11	1361	A	147	614	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT
						NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL
Į.	1					KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI
				<u>'</u>		KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS
						DERVSMGTSSRKPTNSSSSLGALKMSATS\*G
						SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS
12	1262	_	177	13	416	TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI
	}		]			DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM
1	l					VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA
1						ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI
13	1363	_	249	535	105	NLLTMNVY WIEUPHI SPARI IVCDOCTOVISVOPONIIVO
1.3	1202	Α	447	رور	105	WIFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ
						MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS
	·		l			TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI
						DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC
14	1264	_	254	572	201	QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH
						FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT
						*SFFFIFSWGTNGCLLSAITYACYAAICHPLLS
15	1265		257	405	(0)	TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM
						AIEFLLECDQNIT/KLICENT*KNIAKNI*KRRV
				· ·	{	TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI
J.,	1200				401	KQTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK
						SKDVFSKPVNIFWALEESVLGVKARQPKPFFA
[				1	i	AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV
						GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	Α	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW
<u>                                     </u>	1000					VSSLAMKEMLTKTTM
18	1368	A	300	904	1	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL
L					<u></u> i	ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  IHLK*MFYTMSQKMP*PTINLILLLIIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKTASFFK GQPHQSTNALLRRCVR*RYHLS\TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR ISLMSSWDYRRPPO
19	1369	A	302	3	445	NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHILLSHIWKDVQNILLK
20	1370	A	304		1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHQQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW ALVRSRGG
21	1371	A	326	799	1587	GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPO
25	1375	A	384	373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	A	397	383	165	EVKNTNPFIFSGTNLTTWIRST*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLTYKFNTIST KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI LVNKIEDLNKWRNVLLSWIGRRNIINTMT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
28	1378	A	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF YQTFKEEL/II/ILHKLFQTIKYGRILPNSVYETSI TLKPKPEKDL\KENYRPLPLSNIDAK\LNKTLA NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS SWDYRYAPPRPANF\*FLVETGFYYVAQAGL KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK* KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK IFAN
32	1382	A	474	125	471	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KQSK/CQ KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEQPASPSPLKPGTYASET\RPDP HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR APGGSAGS*\CLPSAGGSRGKGWRAAGRQP STR*GRPGRHGGRGE*AGHPEPRQSALQSAG L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA SPQTAAGAGSPVQWALSRATG*TGETGSWC AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP SPASSEVALSSGSCWPDQAPGPARGSPPAPLA PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL SGTS*WRSP*AGTRTQQC*SPWLVPACSSRP L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS*GGRSPAGTGHLGAQTVASPH*GHWPTALSCL WASASPPGPEAPPQTGACIGTNCRYRAASAR RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER GALTHRPRAPDE
34	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRQ KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*QAAFLLRGPGLRPPA/DPSLR/HRN LTELVVAVTDENIVGLFAALLAERRVLLTAS KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH LLDYC*CPPLPRT
36	1386	A	. 512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA FLGLAAGGQTLCPAGELPGHARAQASGAPGS VLIAVPGRRRVHTCGPGPAAPSTRGECPPPAL GHTRPARPRVVPFAPAVPQEPGGQGHGAA/P PATGHSAPRGCPPARAAPTGSATPAPPPAACA AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Į.		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	l	i		peptide	i	/=possible nucleotide deletion, \=possible
<b></b>	ļ	<u> </u>	<b> </b>	sequence		nucleotide insertion
1	}	1				GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD
.	1	ľ			1	GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS
1		1	ļ		1	VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS
1			] .			HGA\GLPGVRVHTQRVHIH*GAG\GCQTPRPR
1		1				LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA
ļ			}	•	İ	ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP
						NHQAVGLEASGALQAGHRDELPTMVQLLDH SPDYPLKGRPHAP
37	1387	A	620	828	1	FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH
3,	1507	^	020	020	1 *	WEASPEMQSKCHQKGKNNQTECFNHVRFLQ
					(	RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE
						FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE
						AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH
						SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH
						WAEFVFSVLVRESKASAVGDDDKVYYFFTE
						RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVO/OGNRGKG
				_		TSEEGKEG*EVPV*LPVSPPLPRPLOKMLDYL
				1		KDKKEVGFFQSIQALMQTC\GEKVMADDEFT
1					,	QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT
						INIIICTVDYLLRLQESI
39	1389	Α	767	1	1030	TLDLTGPLLLGGVPNVPKDFRGRNRQFGGCM
						RNLSVDGKNVDMAGFIANNGTREGCAARRN
						FCDGRRRQNGGTCVNRWNMYLCECPLRFGG
	ļ					KNCEQGEWPASSIPPVTAAWEALLLDVPGTT
						VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL
				İ		RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP
					]	ATVIISVPWYLGLMFRTR\KEDSVLMEATSGG
						PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG
						LRVTDGEWHHLLIELKNVKEDSEMKHLVTM
						TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA
1 .						SEDKVSVRRGFRGCMQVRGGCGGRGEACPS
1						QAAPRL
40	1390	Α	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV
						PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL
				•		RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR
	1201		025		105	VNKSVVLVQVTIM
41	1391	Α .	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD
1	1202		041	<del>-,</del>	416	SSVTNFSIGAK*DILQSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIHWIES
1	-	Ì	.			KASFGDECSHHAYLHDQFWSYWNSLKHRTW
1 1	}	ł		1	Į	QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG
		j		i	-	FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL
43	1202		015	250		RACEPTNIVTL
40	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG
		]		l	ļ	LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP
44	1394		853	452	<u> </u>	QRPMLPPSHAGLARPPPPEPISVP
•••	1374	A	623	434	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK
1		l		l		SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS
		ł	ł	ì		PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG
ļ i			ļ			QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR
45	1305	<del>,                                    </del>	804	270	162	PLTFSTRRNVDPEIPERFR
*	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG
}	.	. 1	Ì			QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT
46	1396	<del></del>	900	1	366	WLSMSMGK
1 10	1370	A	300		366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL
	1	ł	}	ł	1	EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR
		1				VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide sequence	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
47	1397	A	944	162	2	D/KSLAMLPRLVSNSWPQVILPP QLQNLASRGCL*SQLLRRLRRENRLNPGGGG
48	1398	A	963	216	308	CSEIAP\CTPAWVTQRDFFRKKK  HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG AAAAOP**TPRCPAALRAGAHIGRVGRPY
50	1400	A	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE E
	1401	A	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG WEPRGPPSLAPSSP1/TMWTALVLIWIFSLSLS ESHAASNDPRNFVPNKMWKGLVKRNASVET VDNKTSEDVTMAAASPVTLTKGTSAAHLNS MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT AVASSTTAASITTAASSMTVASSAPTTAASST TVASIAPTTAASSMTAASSTPMTLALPAPTST STGRTPSTTATGHPSLSTALAQVPKSSALPRT ATLATLATRAQTVATTANTSSPMSTRPSPSKH MPSDTAASPVPPMRPQAQGPISQVSVDQPVV NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS SGGTKMPATDSCQPSTQGQYMV/DHH*APHP GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL *ELQEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNQVLPNSFGAIEEIC QMLKHI
52	1402	A	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF *SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL VLNAMLSIVPAVVPAGKTRHEKEITCPLIGQE EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV *LMLA*TKDPSIRPTAKELLKHKFIVKNSKKT SYLTELIDRFKRWKAEGHSDDESDSEGSDSES TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ DLVQTLSCLSMIITPAFAELKQQDENNASRNQ AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIQH/CFMITTLKKLGIDGKYLN TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA EAAEHRPEEDRMLSEDPWRPAHMIKGYMPL HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT QKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	A	1044	5	429	SVLTLQTRSPSKPLSRKLMDWEVVSRNSISE DRLETQSRASRSPPVTPNQSQETPVDGKPLAL PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP SSSGIPTPIAVITDALTDLVELILGQPCSEESGR APGTLFLLAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of- peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
57	1407	A	1050	11	430	GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD MH\ISLPMAFLLRTLVRCTSYIIPVTHVLSTPV TCLRRREKDGVIVDVLSDTASNHNGFPVEEH ADDTHPARLQGPTLRSQPMGPLKHKAFEERA NLGLVQRRLRLED
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL LLLAVQQSCLADHLLTASWGGK/DPIPTKALG EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCOPFHLVMRDLLQLGQDIPQGCHY LEENHLIHRDIAARNCLLSCAAPTRAATIGDF GMARYIYRTRYYQLGDRAL/LPRK WMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER ANLMHMMKLSIKVLLQSALSLGRSLDADHA PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF GPLELVEKLCPEASDIATSVRNLPELKTAVGR GRAWLYLALMQKKLADYLKVLIDNKHLLSE FYEPEALMMEEEGMVIVGLLVGLNVLDANL\ CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE HERITDVLDQKNYVEELNRHLSCTVGDLQTK IDGLEKTNSKLQERVSAATDRICSLQEEQQQL REQNELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI HTGEKPYTCGECGKTFRQSANLYAHKKIHTG EKPYTCGDCGKTFRQSANLYAHKKIHTG\ EKPYTCGDCGKTFRQSANLYAHKKIHTG\ YKCKECGKAFKSYYSILKHKRTHTRGMSYEG DEC\QRSLN\RSSILSNHKIIHNEEK\PLKCEKCE KAFNHTSICCRHKKN
64	1414	A	1084	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA SSSQSPEAIKQLLDSGLPSLLVRSLASFCPSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW TALLFLLCHSGSTSGS\HNLG\AQQDQCKISFS FFSWLTTGLITQQRTAIE\NATVAFFLQC\SC HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI S\GFIR\RLFLQLMLEDEK\VTMFLQSPCPLYKG RINATSHVIQHP\MYGAGHKFRTLHLPVSTTL SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415.	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA SVALHKLSNALV
66	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD TLPVAAAFTETVNAYFKGADPSKCIVKITGE MVLSFPAGITRHFANNPSPAALTFRVINFSRLE HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL MTHLK
67	1417	A	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA PYYFLLDLCCSDILRSAICFPFVFNSVKNGST WTYGTLTCKVIAFLGVLSCFHTAFMLFCISVT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RYL
68	1418	A	1106	1	1326	MGKISATGINMGTKCSWALVWHLESYDPKH YEREGMQDWKTASGQSEEATQQSSQKPQPH YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF PGRSRARPPRTRQQRRGAAAGPGRGAVRLG HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA RNSDASGPASLSRTLGRASSPRPPQAPDVTAP SPAALAPRAARGGSRAAALAGAEAEEPLRTL APRPTRAAAPPPPPPPPPLPPGAPPPPVRCVSR RARAPPWRPAATGPPPRPVAPSRKLGSARAP APALQIRKGTSSGLPGRGGGSGPGNNLSSVA GNWRGSSFAVERPGMAKYQGEVQSLKLDDD SVIEGVSDQVLVAVVVSFALIATLVYALFRNV HQNIHPENQELVRVLREQLQTEQDAPAATRQ QFYTDMYCPICLHQASFPVETNCGHLFCGSLT PNSIW
69	1419	A	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	Α	1111	698	23	ALRRLHYVRATKVIFLSFRRPFWREEHIEGGH SNTDRPSRMIFYPPPREGALLLASYTWSDAAA AFAGLSREEALRLALDDVAALHGPVVRQLW DGTGVVKRWAEDQHSQGGFVVQPPALWQT EKDDWTVPYGRIYFAGEHTAYPHGWVETAV KSALRAAIKINSRKGPASDTASPEGHASDMEG QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL QNTTHTRTSH
71	1421	<b>A</b>	1119	2	385	QKQTLQNGYLDSSMDILYLGSLPPELQVSSDE PPGPPEQAGLSQFHLEPETQNPETTEEIQSSLQ QEAAAQLPQLPEVVELSSTKA'EAPALPSQSL EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT GQKVAIKIVNREKLSESVLMKVEREIAILRLI EHPHVLKLHGVYENKKYFPPDELTSGPSMLA QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG GELFDYLVKKGRLTPKEARKFFRQIVSALDFC HSYSICHRDLKPENLLLDEKNNIRIADFGMAS LQVGDSLLETSCGSPHYACPEVIKGEKYDGR RADMWSCGVILFALLVGALPFDDDNLRQLLE KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR LSLEQIQKHPWYLGGNFIS
73	1423	A	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG SDFLCTEWKASNSVPTSVHQLRPADIKVVAA LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG GDGNLETHTTLPNILKKFNPYLLGFSTSTWEG TAGLNVAAEGARARDMPAQAWDLVERMKN SPDINLEKDWKLVTLFIGGNDLCHYCENPEA HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEP\KGSCWPQSALRGQCPGPVLGV TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
76	1426	A	1155	38	410	DDESGQKKLHGLQAILVHEASGTTAITATAT GYQESHLSSAR PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVQYIMNCMLWK KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF
77	1427	A	1162	526	350	T RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV
78	1428	A	1171	1	1293	LRGNLLRVFRQVEKVQEENKWQSPLED MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYQPCDDMDCLSDRCKILQVFDDFIFIFFA MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PPYYQPEEDDEMPFICSLSGDNGIMGCHEIPP LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI CHKVIMGAGKLPRHMRTHTGEKPYMCTICE VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82		A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLEDPRA SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC WGRGHGCGQEALSTSHGYHLFCALLTGFLFA SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF Q
84	1434	A	1192	45 .	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES GACVNQVTVDSITPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSQPDEKLLSALHINSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT GRSPCPSLPGTTRTNSLL
86	1436	Α	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC NLGFFLEGHAVLTCHAGSENSATWDFPLPSC RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		ļ	ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 .		l		peptide	•	/=possible nucleotide deletion, \=possible
[	•			sequence		nucleotide insertion
				1		NADCTWTILAELGDTIALVFIDFQLEDGYDFL
		1		Ì	}	EVTGTEGSSLW
87	1437	A	1216	226	964	GTARFGPMVGFGANRRAGRLPSLVLGVLLV
	]	1			***	VIVVLAFNYWSISSRHVLLQEEVAELQGQVQ
Ì		i		ļ		RTEVARGRLEKRNSDLFAVVGHAOETDRPEG
		ł	1			GRLRPPQQPAAGQRGPREEM\EDDKVKLQNN
					1	ISYQMADIHHLKEQLAELRQEFLRQEDQLQD
l	}	l	1	1	ł	YRKNNTYLVKRLEYESFQCGQQMKELRAQH
1			ł		i	EENIKKLADQFLEEQKQETQKIQSNDGKELDI
ļ			1	ļ	ļ	NNOVVPKNIPKVAENVADKNEEPSSNHIPHG
88	1438	A	1218	1	534	PEFGTTISCGYLMATDVSRRPSVHKAVEIEOE
00	1430	^	1216	*	554	RVKSAGAWIIHPYSDFRFYWDLIMLLIMVGN
			l .		ļ	LIVLPVGITFFKEENSP\PWIVFNVLSDTFFLLD
1		{	ł		ì	
		ľ	ł			LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW
] .						FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR
89	1439		1223	1	743	ALRIVRFTKILSLLRL
ا قوا	1439	A	1223	1	143	MGFDEVFMINLRRRQDRRERMLRALQAQEIE
				ļ		CRLVEAVDGKVGMLTRSNAAPGRHLAMLET
1	•	1				LVVVAPRFVDADNLILNPDTLSLLIAENKTVV
						APMLDSRAAYSNFWCGMTSQGYYKRTPAYI
1			ļ			PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL
[			1			\AFYPPHPDYTWSFDDIIVFAFSCKQ\AEVQMY
						VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
						LEVMVPSSPSSAQSMAVVSADHIGLVISYL .
90	1440	Α	1227	2	349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF
						GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
				'		YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG
<u></u>						LGARLVWFGKLILAKGGHGGISWL
91	1441	A	1245	3	1937	LGSSDVRAPQRSELGAESPSRMVASQAYNLT
						SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD
1		1				VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF
1						QTQAGAQIPIERLASERAITVKVPNNSDWAAR
1 1						GHRSSANSV\VQPQAFVGAVVTLDSSNPAAV
						LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR
!!!				-		PNEHNCSASRRIRPESLQGADHRPYTFFISPGT
1 . 1						RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC
						QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR
į 1						HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML
						TCAVCLVTYMVMAAILHKLDQLDASRGRAIP
						FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM
j l					1	LYGVDSRSGHRHLDGDRAFHRNSLDIFQIATP
					1	HSLGSMWKIRVWHDNKGLSPAWFLQHIIVRD
					,	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
						AASKASFRVPTPS\AALLRFRRLLVAELQRGF
, ,						FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL
)			ļ		}	FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
	1		l			DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV
	-				l	GWGWGPGSTGNGAWASAPCPEPPLSSAAAR
						GKGVHQRLLGKGQHT
92	1442	Α	1246	5	562	VFDEENILNELNDPLREEIVNFNCRKLVATMP
					ļ	LFANADPNFVTAMLSKLRFEVFQPGDYIREG
] ]	ļ			J	}	AVGKKMYFIQHGVAGVITKSSKEMKLTDGS
		1		1		YFGEICLLTKGRRTASVRADTYCRLYSLSVD
						NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
		ì			1	SILLQKFQKDLNTGVFNNQENEILKQIVKH
93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
		ļ				PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN
]	1	j			j	KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT
L 1			1	ł		ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

NO. of neucleotide original per petitide of petitide sequence   No. of neucleotide location   No. of neucleotide   No. of neuronal petitide sequence   No. of neuronal petitide sequence   No. of neuronal petitide   No. of neuronal petit	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Supplement	NO: of	NO: of	hod	ID NO:	beginning		D=Aspartic Acid, E=Glutamic Acid,
Sequence   99496   onespond   seid residue of peptide residue of peptide residue of peptide sequence   Pop	1		l				
1440   1261   3   385   1445   A   1261   3   385   1446   A   1294   1   1456   1456   1446   A   1294   1   1456   1446   A   1294   1   1456   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1448   A   1295   2   2057   1448   A   1295   2   2057   1448   A   1295   2   2057   1448   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1447   A   1295   2   2057   1448   A   1304   118   453   459   459   459   459   450   4							
amino acid residue of sequence peptide s		uence	ļ	1			
residue of peptide sequence   V=Tyrosine, X=Uludnown, **Site peotole, /possible muelodide delicina, \possible muelodide discrition   LpPSGERSEL ARGEMENT   LPSGERS   LPSGERS	uerice	1		1 214 .			
Peptide							
	ł		l			sequence	
1444	1						
94		<del> </del>	├	<del> </del>	sequence		
1444	1						
1444   A   1261   3   385	<b>(</b>		Ì	[		ſ	
1444   A   1261   3   385	1 '			1			
1445   A   1282   2   250   350	94	1444	A	1261	3	385	<del></del>
95 1445 A 1282 2 550 GPRINFYAMDETHIKCLVLRIQEAFGNKLCPA  96 1446 A 1294 1 1456 QLLPPSINEEMEGLATIFACKLAWATELIARI  97 1446 A 1294 1 1456 QLLPPSINEEMEGLATIFACKGPONTORY  98 1447 A 1295 2 2057 QRAMSLASIFTENDESSENDLY PREDICE PREVENCE OF A STANDARD PROPERTY OF		- · · ·					
DPNHFYAMDETHHKCLVLRIQEAFGNKLCPA			ĺ	[			
95 1445 A 1282 2 550 GPRDNPGUDPRTEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSIVPPYITLVVNL VVESTPLANLGRVAQVLRIMRIFILLKARF STGLRSLGATLKYSYKEVGLLLLYLSVGISIPS VVAYTIEKEENGLATTACWWWATVSMTT VQGDVVPGTTAGKLTASACLAR 96 1446 A 1294 1 1456 QLLPPSINENAGLLVGRCLCSAALRPVGDLT SGQVAVRNAPQAGSALAGKPGDNFEIGL YFKKFFDANCNEKDYNPVAAQQGOETEVAP SIVAPVLNKENQCPEGYTCVAVAMA YEEQNQATLEBACQKAAFTQOMLEQLKXQQ EAAQQAATATASEHISRPSAAGRLSSSSEAS KLSSSSAKERNRKKKRKQSGGEEKDED EPQKSSESDSIRRKGFRSIEGNRLTVERKYSS PHOSILSIRGSLFSPRNNSTRISTSFRGRAKDV GESNDFADDEHSTEEDNESRRDSLFVPRRHGE RRNSSLSJTSSSSMLAVPNSKMHSTVDC NGVVSLVGGPSVPTSPVQQLLPEVIIDKPATD DNGTTTETEMRKRKSSSHVSKHFTOSEPPDLP MENASLSJTSSSSMLAVPNNSKOPHLEDPSQR QRAMSLASLITNIVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFILE EVLLKYRITFYNNKOPNULYQHLEKAPVHFM NRYLSYGGPKRPPLVDVLQVALLEFASSKPV CTSPVDDDASSPPGSIPSGTLTSTTGQGALS SELPSTSPSSVAAISSRSVHKFFTOSREPPDLP MHPAPHITEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVYKSSWEELLYRDSFGGYRNGQVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVKSSWEELLYRDSFGGYRNGQVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVKNSSWEELLYRDSFGGYRNGANGVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVKNSSWEELLYRDSFGGYRNGANGVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVKNSSWEELLYRDSFGGYRNGANGVPYRLHAV LVHEQQANAGHTWAYIFDHRESRWMKYNDI AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNN AVKNSSWEELLYRDSFGGYRNGAAVCLMYNNELL SKGLYRDFORDATSFFTLUCHRVEN ARABLEKTLICAGGONNKELL SKGLYRGHDEELSHYNRCCLLILLNDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLLILANDQAAELF ESGEDREVNNCLULDETKEELDSYBLD LAVEDWRAWCSYLGGORMPHLQEKLT DFLYKLLDCSMERSFFEPKLPSYSTHELCER FARMLSLSRTFADOR  98 1448 A 1304 118 453 SGPSSRATHRKEYSONLTSEFTLLQHVEHE  1449 A 1306 3 1660 CCYFCHTTCAPQAPPCPVP	1	:	ĺ			1	
VAPPEIKFEKNALNI.IDLMSIVPEYTIT.VVNI.  VVESTPTLANLGRVAQVVRLMRIFRILKLARH STGLRSLGATIKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACW WAAVTSMIT  96 1446 A 1294 1 1456 OLLPFSNRENAGLLVGRCLCSAALRPVGDLIT SKQVAVVNANPQAGSAKGKFQDNFEFIQ YFKKFDANCNERDYNPVAAGQQGETEVAP SIVAPVLANKPNQCEPGYLAKAGRNPNYGYT SPDTFSWAFLSLFRLMTODYWENLYQLTLRA AETTYMIFLV.IVILLGSLYLVTILLAVW ANA YEEQNQAATAASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRRRKKRKGKSCOMECOLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRRRKKRKGKSCOGGEEKDED EPQKSESEDSIRRKGFRSIEGNRLTYERRYSS PHQSLLSIRGSLFSFRRNSILSFSFRGAKDV GSENDFADDEHSTFEDNESRRDSLFPRRHGE RRNSINLSOTSRSSMLAVFSFGGGEEKDED EQKASSESTENSSSEANSTRANGKMSTVDC NGVVSLVGGFSVPTSPVGQLLPEVILDKPATD DNGTITTETEMKRSSFHVSMPFLEDFOQ RAMSIASILTNITVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLKKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPMALYEIKAFVHFM NRYLSYGSGPKRFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGSIPSANGLHWFTGGGALS SELPSTSPSSVAAISSRSVHKPFTGGALS SELPSTSPSSVAAISSRSVHKPFTGGGALS SELPSTSPSSVAAISSRSVHKPFTGGGALS SELPSTSPSSVAAISRSVHKPFTGGALS VANTKSSWEGLHWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LYHEQQANAGHYWAYIFDHRESRWMXYNDI AVYKSSWEGLEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPTYLEGPSRSDFSK HLKEETIQUITKASHEHBLASPETVLQSAIKLE YARLVKLAQEDTPPETDVRLHHVVVYFIQNQ APKKIIEKTILEGFGORNISASPCLUKHMEVL SGGLRGVHREELISHYRRECLLKLINEQAALEL ESGEDREVNNGLIIMNEFTVPFLPLLLVDEMEE KDILAVEDMRRWCSVLQEMEPHLQEKLT DFLPKLLDCSMEKSPHEPELPSYSTHELCER FARIMISLSSTPADGR  98 1448 A 1304 118 453 SGPSSRATYLFRKEYSQNLTSEPTLLORRVEH LMTCKQGSQRVQGPEDALQKLFEMDHARGV WSQDLLLQVFDOWLOLDIETKEELDSYRLD SIQAMNVALINTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTCAPQAPFCPVPDDLRTALGVHFE	1 .		}			}	
VAPPE,KFFKNALNI,IDLMSNPFTILLVANL	95	1445	A	1282	2	550	GPRDNPG\EDPRFEIVEHFGIAWFTFELVARFA
96 1446 A 1294 1 1456 QILPPSINERACILLY LISVGISTES STGLRS LGATILK YSYKE VILLY LISVGISTES VVAYTIEKEENEGLATIPACWWWATVSMTT VGYGDVYPGTTAGKLTASACILA  96 1446 A 1294 1 1456 QILPPSINERAGILVGRCLSAALREVGDLIT SIGQVAVRIAPQAGSAKAGKGKORDIFETIQ YFKKFFDANGNEDYNYPVAAGQAGGEVEVAP SIVAPVLINKPNOCPEGYICVKAGRINNYGYT SPOTTSWAFLSI FEILMTODYWBRIL YQLTILLA VIVAMA YEEQINQATLEEAGUKEAEFQMILEQLKKQQ EAAQQAATATASEHSREFSAGRISDSSSEAS KLSSKSAKERRINKKKRQKEQSGEEKDED EPOKSESEDSIRKKGFRISPGISTRITYEKRYSS PHOSILISIRGSIFSTRANGISTISFSRGAKDV GISENDFADDEHSTFEDDESRISHSFORTITYEKRYSS PHOSILISIRGSIFSTRANGIFSPSAGRISDSOSEAS KLSSKSAKERRINKKKRQKEQSGEEKDED EPOKSESEDISTRIKGFRISPGISTPSTUTYERRYSS PHOSILISIRGSIFSTRANGIFSPSAGRAKDV GISENDFADDEHSTFEDDESROLIFYERRYSS PHOSILISIRGSIFSTRANGIFSPSAGRAKDV GISENDFADDEHSTFEDDESROLIFYERRYSS PHOSILISIRGSIFSTRANGIFSPSAGRAKDV GISENDFADDEHSTFEDDESROLIFYERRYGS WORKINGTITYERRYSS OUR KIEGNEVYRE PHOSILISIRGSIFSTRANGIFSPSAGRAKDV GISENDFADDEHSTFEDDESROLIFYER FUR PHOSILISPATION OF THE POLITICAL			}				
STGLRSLGATLKYSYEVGULLLY.SVGISITS VVAYTIEKERNEGLATIPACWWWATVSMTT VGYGDVVPGTTAGKLTASACILA  96 1446 A 1294 1 1456 QILPPSNRENAGILVGRCLCSAALRPVGDLT SIQUAVRANAPQAGSAKAGKGKPQDNFEFIQ YFKKFPDANCNEGDYNPVAAGQGQETEVAP SIVAPVANKPNOCEGEGYLAKGRNPNYGYT SPDTTSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIFILVI.VULLOSLYLVTLLAVVANA YERONQATLEAEQKEAETQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSRSAKERRRRKKKGKCSGGEEKDED EPOKSSESDSIRRKGRRRSFSKGVGSGGEEKDED EPOKSSESDSIRRKGRRSPSIGGRRLTYERRYSS PHOGLISIRGSLFSPRRNSIFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSBMLAVFPANGKMHSTVDC NGVYSLVGGFSVTSPYOQLIPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLKGGNCVCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFM NRYLSYGSGFKRFLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSFSQTLPSTTEQQGALS SELPSTSPSVAAISRSVHFTGGSRIPPDLP MHPAPRHTEEGLSVLESCLHRWRTGENDTR DLOESISRHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDIRESRWMXYNDI AVTKSSWEELVRDSFGGFRNSASVALMYIN DKAQFLIQEDLIKTGQLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQKLLAS QKLRSESTSVTAQAAGDFKYLEOPSRSDFSK HLKEETIQITTKASHEHEDXSPETVLQSAIKLE YARLVKLAQEDTPFETDYRLHHVVYYFIQNQ AFKKIEKTLLEQFGDRNLSTDERCHNIMKVA QAKLEMIRPEVNLEEYEEWHQDVKFRETT MYLIGIGLBNFQRSSYDGLJFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLLLVDEMEE KDILAVEDMRNRWCSYLQQEMEPHLQEKLT DFLPKILDCSNEIKSTHEIPPKLPSYSTHELCER FARIMLSLSKTPADGR  98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPILLGHRVEH LMTCKQGSGRVQGPEDAJ CKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTCATQAPPCFVPPDLLRTAGCHFHE			l				
96 1446 A 1294 1 1456 QILPPSNRENAGILVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAKGKFONFERIQ YFRKFFDANKONERDYNPVAAGQQGETEVAP SIVAPVINKPNQCPEGYICVKAGRNPNYGYT SFOTTSWARLSIFRILMTQDWENLYQLTLRA AETTYMIFALVI.VILLGSLYLVTILLAVVAMA YEEQNQATLEEABQKEAEFQQMLEQLKKQQ EAAQQAATATASEHRREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRQKEQSGEEKDED EPQKSSESDSSTRKGFRISGSIFSTRRISTIFSRGRALDV GSENDFADDEHSTFEDNESTBLFVPRRHGE RNISNLSQTSRSSSRMLAVFFANGKMHSTVDC NGVVSLVGGSVPTSPVGQLLPEVIUDKPATD DNGTITTETMRKRRSSSHVSMPFLEDPSQR QRAMSIASILTNITVE  1447 A 1295 2 2057 IQTQLFYTSSQQLRKGGNCVCKKMQMNFIAE EVILKYRITFYNNNKGPMLYIEIKFVHFM NRYLSYGSGYKRPLVDVUQVALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELISVLESCLHWRTEIENDTR DLQESISRHRTIELMYSDKSMQVYRLHAV LVHEQQANAGHYWAYIDFIRESRWMXYNDI AVTKSSWELVRDSFGGYRNASAYCHMYIN DKAQFILQBOLIKTQQLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLLQPSKSKSK HKREETQIITKASHBEHDSKSPETVLQSAIKLE YARLVKLAQEDIPPFTDYRLHHVVYFIQNQ APKKIEKTLLEGFGDRNLSFDERCHNIMKVA QAKLEMKPEEVNLEYSEWHQDVKKRETT MYLDGLENFQRSSYDGSLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEPREVNOGLIMMEFTVFFIFLLLVDEMEE KDILAVEDMRNRWCSYLGGMEPHLQEKLT DFLPKILDCSMEKSFHEPFKLPSYSTHELCER FARIMLSLSSTPADGR 98 1448 A 1304 118 453 SQPSSRATYLIRKEVSQNLTSEPTLLQHVPEH LMTCKQGSQRVQGPEDALQKLFFEMDAHGRV WSQDLILQVRDGWLQDLIDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTCAPQAPPCPVPPDLLRTALGVHPE	}						
96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGGGETEVAP SIVAPVLNKPNQCPEGYICVKAGRIFNYGYT SFDTFSWAFLSJFRIMTQDYWENLYQLTLRA AEITYMIFLVA.VILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEISFSAGGRLSDSSSEAS KLSSKSAKERRNRKKKQKGKGGGEKDED EFGKESEDSIRRKGFRSIEGNRLTYEKRYSS PHOSLISIRGSLFSPRRNSRISLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRIMLAVFPANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHPMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSFVDDIDASSPPGSSDTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTGSRIPPDLP MHPAPRHITEEELSVLESCHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFKELEEWDAQLAQKALQEKLLAS QKLRESSTSVITAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSFETVLQSAIKLE YARLVKLAQEDTPETIPTJHHVVVYIGNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEVNLEEYEEWHQDYRKPRETT MYLLIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKINEQAAELF ESGEDREVNNGLIMRFVPFIPFLLLLVDEMGE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPFPFIPLLLVBEMEE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPFPFILLLVPBEMEE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPSTHELCER FARIMLSLSRTPADGR	}		}				
96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLIT SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYNPVAAGGGETEVAP SIVAPVLNKPNQCPEGYICVKAGRIFNYGYT SFDTFSWAFLSJFRIMTQDYWENLYQLTLRA AEITYMIFLVA.VILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEISFSAGGRLSDSSSEAS KLSSKSAKERRNRKKKQKGKGGGEKDED EFGKESEDSIRRKGFRSIEGNRLTYEKRYSS PHOSLISIRGSLFSPRRNSRISLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRIMLAVFPANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHPMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSFVDDIDASSPPGSSDTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTGSRIPPDLP MHPAPRHITEEELSVLESCHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFKELEEWDAQLAQKALQEKLLAS QKLRESSTSVITAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSFETVLQSAIKLE YARLVKLAQEDTPETIPTJHHVVVYIGNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEVNLEEYEEWHQDYRKPRETT MYLLIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKINEQAAELF ESGEDREVNNGLIMRFVPFIPFLLLLVDEMGE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPFPFIPLLLVBEMEE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPFPFILLLVPBEMEE KDILAVEDMRRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKPSTHELCER FARIMLSLSRTPADGR	L				ļ!		
SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ YFKKFFDANCNEKDYPVAAGGGGETEVAP SIVAPVINKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIFLVAVIULGSLYLVTILLAVVAMA YEGONQATLEBAEGKEAFFQMLEQLKKQQ EAAQQAATATASEHSPAAGRLSDSSSEAS KLSSKSAKERNRKKRKQKEQSGGEEKDED EFÇKSESEDSIRRKGFRSIEGNRITYEKRYSS PHOSLLSIRGSLFSPRRNSITSLFSFRGARADV GSENDFADDEHSITEDNESRRDSLFVPRRHGE RNSNLSQTSRSSRMLAVPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILKYKITYNNNKGROMLYLEKAFVHFMI NRYLSYGGGPKRPPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSISSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLIRWRTEIENDTR DLOESISRIRHTIELMYSKMIQVPYRLHAV LVHEQQANAGHWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKBLEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDRYLEQPSRSDFSK HLKEETIQUITKASHBHEDKSPETVLQSAIKLE YARLVKLAQEDTPETPLTLICHTQVHHVVVYIPONQ APKKIIEKTLLLYGAAGDRYLEPTSPSDFSK HLKEETIQUITKASHBHEDKSPETVLQSAIKLE YARLVKLAQEDTPETPLTLICHTQVHHVVVYIPONQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEGVNAGLINGFLICHCAYONKELL SKGLYGGHDEELISHYRRECLLKLNEQAAELF ESGEDRESVNNGLINGFPOPPPIDLLICHQVNKKELL SKGLYGGHDEELISHYRRECLLKLNEQAAELF ESGEDRESVNNGLINGFPOPPPIDLLUPDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSBEKSFHEPFKLPSYSTHELCER FARIMLSISTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGGQRVQGPBQAGKLGLIFENDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPYPPDLLLATAGVHFE	96	1446	Α	1294	1	1456	
SIVAPVILNKPNOCPEGYICVAGRNPNYGYT SFDTFSWAFLSIFRIMTODYWENLYQLTIRA AETTYMIFL/JL/JUGSLYILVTILIAVVAMA YEQONQATILEAEQKEAEFQOMLEQLKKQQ EAAQQAATATASERPSAAGRISDSSSEAS KLSSKSAKERRNRKKRKQKEQSGGEEKDED EFQKSESEDSIRRKGFRSIEGIGNILTYEKRYSS PHOSILSIRGSLIFSPRINSRTSLISFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVTFSVVGQILFEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTITVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLIKYRITFYNNKGPMILYIEIKAFVHFMI NRYLSYGSGPKRFTDVLQVALLEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISRSVHKPFTQSRIPPDLP MPPAPHITEEELSVLESCLHRWRTEIENDTR DLQESISRHRTTELMYSDKSMQVPYRLHAV LVHEQQANAGHTYWAYTFDHRESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFFKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITIKASHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTILLEQFGDRNLSFDERCHNIMKVA QKALEMIKPEEVNLEEYEEWHQDYKFRETT MYLIGLENFQRESYDSULJELICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLINEQAALIF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLFKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 GGFSSRAYTLHRKEYSQNLTESPTILQHRVEH LMTCRQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLIDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			i				
SFDITSWAFLSLIFRLMTODYWENLYQLTIRA AETTYMIFLV/LVILLGSLYLVTILLAV/VAMA YEEQNQATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRINKRKOKQEGGEMEDL EFQKSESEDSIRKKGRKFRSIEGINRLTYEKRYSS PHQSLLSIRGSLFSPRRINSTISFSRGRAKDV GSENDFADDEHSTFEDNESRDSLFYPRRHGE RRNSNLSQTSRSSSMLAVFPANGKMHSTVDC NGVVSLVGGPSVFDVGQLJEPVIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  1447 A 1295 2 2057 IQTQLPYKSSQQLFXGGNCVRCKMQMNFIAE EVILKYRITFYNNKGPNMLYEIKAPVHPMI NRYLSYGSGPKRFDVJLQYALEFASSKPV CTSPVDDDASSPPSGSIPSQTLPSTITEQQGALS SELPSTSPSSVAAISSRSHIKFTQSRPPDLP MIPPAPRHITEELSJESCLHRWRTEIENDTR DLQESISRHRTTELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVARDSGGYRNASAYCLMYIN DKAQFLQEDDLIKTGQPLVGGTLPPDLRDFV EEDNQRFKKLEEWDAQLAQKALQKLLAS QKLRSSETSVTTAQAAGDPKYLEQPSRSDFSK HIKEETIQIITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDTPPFDTVRLHHVVVYTQNQ APKKIEKTLLEOFGGRNLSFDERCHNIMKVA QKALEMIKPEEVNLEYEEWHQDYKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLINEQAAELF ESGEDREVNNGLIMNEFIVPFPLLLLYDEMEE KDILAVEDMRNRWCSYLGGEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 GGFSSRAYLHRKEYSQNLTSTEPTLLQHIRVEH LMTCRQGSQRVQGPEDALQKLFEMDAHGRV WSQDLLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPYPPDLLRTALGVHPE	}						
AETTYMIFILVILIGISTIVITILIAVIVAMA YEEQNQATLEAEQQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKKRQKEQSGGEEKDED EFQKSSESDSIRKGFRISTIEGNRITYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSILSGTSRSSRMLAVFFANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNIVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNVRCKMQMNFIAE EVLLKVRITFJANNKGPMLYTEIKAPVHPM NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDIDIASSPPSGSIPSQTLPSTTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTGSRPPDLS MFPAPRHITEELSVLESCHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDIRESRWMKYNDI AAVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAKKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVVILEFYEEWHQDYRKFRETT MYLIGLBIFFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLINEQAAELF ESGEDREVNNGLIIMNEFIVPFIPLLLVDEMEE KDILAVEDMRNRWCSYLOGEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCQGSQRVQFEDALQKLFEMDAHGRV WSQDLLLQVRDWGWCJLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT
YEEQNQAATASEHSEQKEAEFQQMLEQLKKQQ EAAQQAATASEHSREPSAAGRISDSSSEAS KLSSKSAKERRNRKKKRÇKÇQKGGEEKDED EFQKSESEDSIRRKGFRFSIEGNBLTVEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTIEDNESRRDSLFVPRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVQQLLPEVIDDKPATD DNGTTTETEMKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYKLHAV LVHEQQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRSESTSVTLAAGAGDPKVLQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQAAGDPKVLQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQAAGDPKVLQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE SKGLVRGHDEELISHYRRECLLKINEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLQQEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQQSQRVQGPEDALQKLFEMDAHGRV WSQDLLLQVRDGWLQLLDIETKEELDSYRLD SIQAMVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA
EAAQQAATATASEHSREPSAÁGRLSDSSSÉÁS KLSSKSAKERNRRKKRKQKEQSGGEKDED EFQKSESEDSIRRGFRFSIEGIRLITYEKRYSS PHQSLLSIRGSLFSPRRNSTSLFSERGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASLITHTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGGGFKFPLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHR WRTEIENDTR DLQESISRHRTTELMYSDKSMIQVPYKLHAV LVHEQQANAGHYWAYIFDHRESR WMK YNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTIAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKILLEQFGDRNLSTDERCHNIMKVA QAKLEMIKPEVNLEEYEEWHQDYRKFRETT MYLIIGLBNFQRESYIDSLIFICAYQNNKELL SKGLYRGHDEELISHYRRECLIKLINEQAAELF ESGEDREVNNGLIINMEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKCQGSQRVQGPEDALQKLFEMDAHIGV WSQDLLQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	{						AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA
KLSSKSAKERRNRKKRKQKEQSGGEKDED EPQKSESEDSIRRKGFRPSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSFMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATID DNGTITETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSGTLPSTTEQQGALS SSLPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTEEJSVLSSCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIPPHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKIKRESETSVTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTILEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLIFLICAYQNNIKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 GGPSSRAIYLHRKEYSQNLTSEPTILLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTICAPQAPPCPVPPDLLRTALGVHPE							YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ
EFÇKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATID DNGTITTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVYLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSOSIPSQTLPSTTEQQGALS SELPSTSPSSVAISSRSVHKPFTQSRPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYTFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETTVQSAKLLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEETYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFILCAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNETIVPFILLVDEMEE KDILAYEDMRRIWCSYLGGEMEPHLOEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMIVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	J .						
PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMKRRSSSFHVSMDFLEDPSQR QRAMSLASTINTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYTEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HIKREETIQITKASHFIEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLEIN-QRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFIPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SQFSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFENDAHGRV WSQDLLLQVRDGWUQLLDDIETKELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
GSENDFADDEHSTFEDNESRRDSLFYPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNITVE  97 1447 A 1295 2 2057 IQTQLFTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLFSTTEQQALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPHITEELELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPTETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLEINFQRESVIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNFEIVPFLLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SQFSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLLQVRDGWLQLLDDIETKELDSYRLD SQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLIPEVIDLKPATID DNGTTTITETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE  97 1447 A 1295 2 2057 IQTQLPTKSSQQLKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQVALEFASSKFV CTSPVDDIDASSPPSGSBPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIPDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKITLLEGFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVJEWHQDYKFPETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLFKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAYLHRKEYSQNLTSEPTTLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	i						
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE 6VLLKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSFVDDIDASSPPSGSIPSQTLPSTITEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPARRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTQQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAHKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKITLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGGLENFQRESYIDSLLFILCAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFIPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARMLSSRTPADGR  98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLLQVRQBWQLLLDUETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
97 1447 A 1295 2 2057 IQTQLPTKSSSGLYKMQMNFIAE  97 1447 A 1295 2 10TQLPTKSSQQLRKGGNCVRCKMQMNFIAE  EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI  NRYLSYGSGPKRFPLVDVLQYALEFASSKPV  CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS  SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP  MHPAPRHITEELSVLESCLHRWRTEIENDTR  DLQESISRHRTIELMYSDKSMQVPYRLHAV  LVHEQANAGHYWAYIFDHRESRWMKYNDI  AVTKSSWEELVRDSFGGYRNASAYCLMYIN  DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV  EEDNQRFEKELEEWDAQLAQKALQEKLLAS  QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK  HLKEETIQIITKASHEHEDKSPETVLQSAIKLE  YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ  APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA  QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT  MYLIGLENFQRESYTDSLLFLICAYQNNKELL  SKGLYRGHDEELISHYRRECLLKLINEQAAELF  ESGEDREVNNGLIIMNEFIVPFILLLVDEMEE  KDILAVEDMRNRWCSYLGQEMEPHLQEKLT  DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER  FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRATYLHRKEYSQNLTSEPTLLQHRVEH  LMTCKQGSQRVQGPEDALQKLFFMDAHGRV  WSQDLILQVRDGWLQLLDIETKEELDSYRLD  SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPFCPVPPDLLRTALGVHPE							
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAGHIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFJQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYYLIIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQYRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1 . 3						
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVILLKYRITFYNNNKGPNMLYIEIKAPVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTIEELSVLESCLHRWRTEIENDTR DLQESISRIHRTEIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKILDCSMEIKSFHEPPKLPSYSTHELCER FARIMI.SLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESTISVITAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	07	1447		1006		0065	
NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTTELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	9'	1447	A	1295		2057	
CTSPVDDIDASSPPSGIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHIKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					·	:	
SELPSTSPSSVAAISSRSVIĤKPFTQSRÎPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		•					
MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQYRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					į		
DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE							
LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\(\text{DLIKTGQPLVGIETLPPDLRDFV}\) EEDNQRFEKELEE\(\text{DAGDPKYLEQPSRSDFSK}\) HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					ļ		
AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	j	·					
DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	[	•	ĺ		j		
EEDNQRFÉKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HILKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTILLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					ļ		
QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE					j		
HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	]				j		
YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR 98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	[ [		ĺ	ĺ	j		
APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE				1		{	
QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			]				
MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	] ]	j	J		J		
SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE						ĺ	
ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			- 1			ľ	
KDILAVEDMRNRWCSYLGQEMEPHLQEKLT DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER FARIMLSLSRTPADGR  98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS  99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	1 1	ł	- 1	ł	1	ļ	
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	] ]	J		Į	·	ļ	
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	]	j		[			
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSQNLTSEPTLLQHRVEH LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		-				. ]	
LMTCKQGSQRVQGPEDALQKLFEMDAHGRV WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	98	1448	A	1304	118	453	
WSQDLILQVRDGWLQLLDIETKEELDSYRLD SIQAMNVALNTCSYNSILS 99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE			]				LMTCKOGSORVOGPEDALOKLFEMDAHGRV
99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE	]	1	1	ĺ		ł	
99 1449 A 1306 3 1660 CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE		.[	ĺ			1	
	99	1449	A	1306	3	1660	
				_	1		

NO. of collection concepted   No. of collection concepted   No. of collection   No.	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide							
Sequence				1 .		location	
100				USSN		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino acid residue of peptide residue of peptide sequence   T=Threonine, V=Valine, W=Typtophan, Y=Typtorine, X=Unatowa, Subported, Sequence   Sequence	seq-	uence		09/496	correspondi	to last amino	
Peptide   Sequence				914	ng to first		
peptide   sequence				}	amino acid		
		[	١.	· '			
		ļ i			peptide		
SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP					sequence		
TTCTVILLAESEGERERWIQVI.GELQRILLU    APPRPRPYTHEA/DTOIGLIPHICAAILD    APPRPRPYTHEA/DTOIGLIPHICAAILD    APPRPRPYTHEA/DTOIGLIPHICAAILD    APPRPRPYTHEA/DTOIGLIPHICAAILD    APPRPRPYTHEA/DTOIGLIPHICAAILD    APPRPRPYTHEA/DTOIGLIPHICAAILD    EVEVEKIPESGCQVLAAGGFALYPLINEAAPLAL    GAGLUPELPESGGCI.GEALGAAPTVLOVA    SIGLIGBRILCVGAAGGFALYPLINEAAPLAL    GAGLUPELPESGGCI.GEALGAPTVVQVAV   SIGLIGBRILCVGAAGGFALYPLINEAAPLAL    LITTAGIYYDGAGRSSRCHELLWPAAPPNOW   SIGLIGBRILCVGAAGGFALYPLINEAAPLAL    LITTAGIYYDGAGRSSRCHELLWPAAPPNOW   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLRINGLAE   KKWPLNPEGGLELYGTKVRLTYLEVCKWLL   POSKUGTVYTAAVAGVVLVVXLLNDKLQQ   QLPMPIGELLTIGATGISYGMGLKHRFEAD   POVAPNTQLFSKLUGSAFILAVOGFALISLGK   FALRHGYRVDSNQWWMRDV   POVAPNTQLFSKLUGSAFILAVOGFALISLGK   FALRHGYRVDSNQWWMRDV   SHOMPARMSTER     1452 A 1363 542 2 AHLIMINIALTYDLLYVITSPFLHYYASGEN   HOPPARMSTERS     165							
ARPRPRPVTILKEAYDNOLPHILCHPHICAGE   ARPRIT   AR			1				
	1 1	[		ĺ			
							I
VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ   SLGLLGDBLCQAAGGGALJPLBAAPIAL   GAGL-VPEL-PPSRGGLGEALGAVELSISFIL   LFTTAGIYVOLAGKESRGHELLWPAAPMGW   GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL   KKWRPLWPGSIELT-YGTEKVRLTYLRNQLAE   KDEFDIPDLITDINSRRQLFRTKSKRFFFTVSE   EQQKQQRREMLDPFVRSKLISPFTHFINILV   HVOPANGRRGABDKSP   EQQKQQRREMLDPFVRSKLISPFTHFINILV   HVOPANGRRGABDKSP   EQQKQQRREMLBDFVRSKLISPFTHFINILV   HVOPANGRRGABDKSP   SLCVFGPVDTGTFAVMSVMVGSVTESLAPQA   LNDSMINSTABAARVQVASTLSVLVGLFQV   GLGLHHFGFVVTTLSEPL-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTTAAAVQVF   VSQLKVVFGLHLSSHSQDFLSIK-VRGYTVAAVAGVF   VSSHGNHALQEVLRSGHGTEFTKQKHLEAAD   HGHPPARMSIPSR   HGHPPARMSIPSR   HALLMILLALUTDLLVLTSLPFLHYYASGEN   WFGDFMCKFRFSHFHVLSSHFLTCSSFFRY   CVTHPMSGCSIH-TRCAVVACAVVWISLVA   VPMFFLITSTHRTNSSACLLDLTSDBELNTIK   VYNLLILTAALLCLPLVIVTLCYTTIHTLTHGHAN   USCLKQKARKTILLL   GVGLYVDCSDSANVYATIR-VLSLQCHHEL   GGDLLHYSFTVLAVIGPDSTHRAATTAALLSF   FLVPMLLEQ   GGDLLHYSFTVLAVIGPDSTHRAATTAALLSF   FLVPMLLEQ   GGDLLHYSFTVLAVIGPDSTHRAATTAALLSF   FLVPMLLEQ   GGDLLHYSFTVLAVIGPDSTHRAATTAALLSF   FLVPMLLEQ   GGGLLYLFAAVEGVMRVTGGSNHLAVVLD   GRGGHEIRFAFYFGGRICVG   KSREERCNENTFCPPVFF   GGGLLYLFAAVEGVMRVTGGSNHLAVVLD   GRGGHEIRFAFYFFGFILLSFTRAK   CSGWMEREWYDDAFWSFLASLLLIVIMFLW   RPSA   GGGLLYLFAAVEGVMRVTGGSNHLAVVLD   GRGGHEIRFAFYFFGFILLSFTRAK   CSGWMEREWYDDAFWSFLASLLLIVIMFLW   GRGGHEIRFAFYFFGFILLSFTRAK   CSGWMEREWYDDAFWSFLASLLLIVIMFLW   GRGGHEIRFAFYFFGFILLSFTRAK   CSGWMEREWYDDAFWSFLAGGRICVG GCPCPVSSLHFILGORPHCCAFGCGLC   FEICPPGGSPPFL VNCKMTSGFFWTCRTDSRVF   QNAGSHEIRFAFYFFGGFILVGENGGL   FEICPPGGSPPFL VNCKMTSGFFWTCRTDSRVF   QNAGSHEIRFAFYFFGGFILVGWERGSGL   FEICPPGGSPPFL VNCKMTSGFFWTCRTDSRVF   QNAGSHEIRFAFWYGGGCGFLGDWLGTKYF   VANNUSTSGFGGAAVGCCVGCFGCGCGCC   GNASCLCSGVFTAHRSLDLVGSN   YFALSLCAALFESDVTGTTDSRVF   VANNUSTSGFGGGAVGCCVGCFGGGGCC   GNASCLCSGV	ľ			1			
SIGLIGBRICYGAAGGRAITYILINFAAPIAL						l	
GAGL-VPEL-PSSRGGLGALGAVELS.ISFILL   LFTTAGIYVOLAGKKSRGHELL WPAAPMGW GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL   KKWURPLNPEGSLF1-KJGTEKVRLTYLRNJLAE   KDEFDIPDLIDNSRQLFRTKSKRFFFTVSE   EQQKQRREMLEDFFVRSKLISPFTFNIHLV   HVQPANGRFGARDKSP   EQQKQRREMLEDFFVRSKLISPFTFNIHLV   HVQPANGRFGARDKSP   EQQKQQRREMLEDFFVRSKLISPFTFNIHLV   HVQPANGRFGARDKSP   EQQKQQRREMLEDFFVRSKLISPFTFNIHLV   HVQPANGRFGARDKSP   SLCVFGPVDTGTFAVMSVMVGSVTESLAPQA   LNDSMINSTADAARVQVASTLSVLVGLFQV   GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF   VSQLEY-VFGLHLSSRGPLSLYTVLEVCWKL   PQSKVGTVVTAAVAGVVLVVYKLLNDKLQQ   QLPMFPGELLTLIGATGISYGMGLKHRFFAG   PPVAPNTQLFSKLVGSAFTIAVYLCWKL   PQSKVGTVVTAAVAGVVLVVKLLNDKLQQ   QLPMFPGELLTLIGATGISYGMGLKHRFFAG   PPVAPNTQLFSKLVGSAFTIAVYLGKHLEAAD   HGHPPARMSIPSR   HGHPPARMSIPSR   SHGNHALQEVLPRSGHGTEFTKQKHLEAAD   HGHPPARMSIPSR   WIFGDFMCKFRFSFHFNLYSSLFLTCTSFHY   CVIHFMSCFSHFKTRCAVVACAVVWIISLVA   VVPMTFLITSTNSTNSRACLDLTSSDELNTIKW   VILLITAALCLPLVIVTLCYTTHTLTHGHAA   USCLKGKARRITILLL   GTGLLYVTLCYTTHTLTHGHAAN   USCLKGKARRITILLL   GTGLLYVTLTAVAGPSTNTKW   VILLITAALCLPLVIVTLCYTTHTLTHGHAAN   USCLKGKARRITILLL   GTGLLYVTLTAVAGPSTNTKATATAALLSF   FLVPMLLEQ   GLGLHYSFTVLAVIGPSTNTATTAALSF   FLVPMLLEQ   SNSWEDRSSMSLWTQNITVCPVRNVTROGG   FGPWSFWQPCEHLDGDNSGSCLCRARSCOSP   RPRCGGLDCLQPAHHANCSRNGAWTPWSSW   ALCSTSCGIGPQVRQRSCSNPARHGGRICVG   KSREERFCNENTTCPTVPF   GLGLLYLTFAAVGSVMRVTGGSSNHLAVVLDD   III.AVIDSIPVWFIFISLAQTMKTLRIKNTVKF   SLYFIFRNTLTAVLASIVPGAGENGCL   GEGULYLTFAAVGSVMRVTGGSSNHLAVVLDD   III.AVIDSIPVWFIFISLAQTMKTLRIKNTVKF   SLYFIFRNTLTAVLASIVPGAGERQSCL   FEIGPQGSPPFLVNCKMTSGTFWTCRTDSRVF   QNANFSNAAHSEDQPTF   GGGLYPTSSLFTLIVIMFLW   QRGGHEIRFAFYPFGHPLLSPQICLAFETPRG   CPPVSSLHTSISQRR PPVCGERQSCL   FEIGPQGSPPFLVNCKMTSGTFWTCRTDSRVF   QNANFSNAAHSEDQPTF   VNCKMTSGTFWTCRTDSRVF   QNANFSNAAHSEDQPTF   VNCKMTSGTFWTCRTDSRVF   QNANFSNAAHSEDQPTF   STSPATSHSVAQAECSGVTTAHRSLDLVGSSN   PFFVTRSFSVAQAECSGVTTAHRSLDLVGSSN   PFFVTRSFSVAQAECSGVTTAHRSLDLVGSSN   PFFVTRSFSVAQAECSGVTTAHRSLDLVGSSN   PPLSLICAAVITSCOGGFLGDWLGTRK   VNVLSTSGGGGAAVGCCVGGGGGGCGCGCGCGCGCGCGCGCGCGCGC		1					
LETTAGIYVDGAGRKSRGHELLWPAAPMGW   GYAAPYLIVPESINDIVPDWSEWQTVPL    KKWPEPLNPEGSLFLYGTEKVRLTYLRNQLAE     KDEFDIPDLITDNSRRQLFRTKSKRAFFFRVSE     EQQKQQRREMLKDPFYNSKLSISPTIVFINILV     HVOPANGRGARDKSP     SLCVYGFVDTGTFAVMSWMGSVTESLAPQA     LNDSMINSTARDAARVQVASTLSVLVGLFQV     GLGLHHGFVYTLSEPLNRGYTTAAAVQVV     VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL     POSKVGTVYTAAVAQVVLVYKLLNDKLQQ     QLPMPIPGELLTLIGATGISYGMGLKHRFEAG     PPVARNTQLFSKLVGSAFTIAVVGFAIAISLGK     FALRHGYRVDSNQVWMMDI     101						ł	l
GYAAPYLTVESENSUPDVRRAEWVQTVPL   KKWRPLNPEGSILI VGTEVKITYLROLAE							1
KKVVRPLNPEGSLEL YGTEKVRLTYRIOLAE					1		·
							1
EQQKQQREMIKDPFVRSKLISPTINFHLY							l
HVÖPÄNÖRPGARDKSP		l					
100							`` ``
LNDSMINETARDAARVQVASTLSVLVGLFQV   GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF   VSQLKYVFGLHLSSRSPLSLITYTLEVCWKL   PQSKVGTVVTAAVAGVVLVVVXLLNDKLQQ   QLPMPIFGELTLIGATGISYGMGLKRIFEAGI   PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK   IFALRHGYRVDSNQVWVMPDV	100	1450	_	1210	018	100	
	100	1430	^	1219	310	130	
VSQLKYVFGLHLSSHSGPLSLIVTVLEVCWKL   PQKVGTVVTAAVAGVVLVVVKLLNDKLQQ   QLPMPIPGELLTLIGATGISYGMGLKHRFEAG   PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK   FALRHGYRVDSNQWMRDV     101							
PQSKVGTVVTAAVAGVVLVVVKLLINDKLQQ   QLPMPIPGELTLIGATGISYGMGLKHRFAGI   QLPMPIPGELTLIGATGISYGMGLKHRFAGI   PPVAFNTQLFSKLVGSAFTIAVVGFAIAISLGK   FALRHGYRVDSNQVWVMRDV							
OLPMPIPGELTLIGATGISYGMGLKHRFEAĞI   PPVAPNTQLFSKLVGSAPTIAVVGFAIAISLGK   FFALRHGYRVDSNQVWVMRDV							
PPVAPNTOLFSKLVGSAFTIAVVGFAIAISLGK   IFALRHGYRVDSNQVWMRDV			l				1
1451			1				1 3
101							· ·
SSHGNHALQEVLPRSGHĞTEFTKQKHLEAAD   HGHPPARMSIFSR     102   1452   A   1363   542   2   AHLIMINIALITDILLIYLTSLPFLIHYYASGEN   WIFGDFMCKFIRFSFHNLYSSILFLTCFSIFRY   CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA   VIPMTELITSTNRTNRSACLDLTSSDELNTIKW   YNLILTAILICLPLVIVTLCYTTIIHTLTHGHAN   DSCLKQKARRITILLI     103   1453   A   1371   2   410   CHSTESSDFILFQDYLLGGLCPLHSGCLQVIC   SFNEHGYHLFQAMRLAVEEINNSTALLPNITL   GYQLYDVCSDSANVYATLRVLSLPGQHHEL   QGDLHYSPTVAVIGPDSTNRAATTAALLSP   FLVPMLEQ     104   1454   A   1376   3   432   NSRVEDRSNIMSLWTQNITVCPVRNVTRDGG   FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP   RPRCGGLDCLGPAHHANCSRNGAWTPWSW   ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG   KSREERRCNENTPCPYPIF     105   1455   A   1379   2   396   GLGLYLIFAAVEGVMRVIGGSNHLAVVLDD   IILAVIDSIFVWFIFISLAQTMKTLRRKNTVKF   SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK   CQSDWMERWVDDAFWSFLPISLILIVIMFLW   RPSA     106   1456   A   1383   1   432   EDGHGGWSSRCLVDHAEGHREPWKRLCTW   QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG   CPPVSSLHFISLQRLPRDCQELFQVGERQSGL   FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF   QNANPSNAAHSEDQPTP     107   1457   A   1386   719   558   FFFVTRSHSVAQAECSGVTAHRSLDLVGSSN   YPALSLQSSWDHRHTWLIFAFL   TVANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	101	1451	A	1353	220	445	
HGHPPARMSIFSR		•				1	
1452							
WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY	102	1452	A	1363	542	2	
CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA VIPMTFLITSTINTRINRSACLDLTSSDELNTIKW YNLILTALLCLPLVIVTILCYTTIHTTTHGHAN USCLKQKARRLTILLL  103 1453 A 1371 2 410 CHSTESSSDFILPGDYLLGGLCPLHSGCLQVVC SFNEHGYHLFQDYLLGGLCPLHSGCLQVVC SFNEHGYHLFQDAWRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ 104 1454 A 1376 3 432 NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGFWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAHHANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF 105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IIILAVIDSIFVWFIFISLAQTMKTLRIKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLFSLILIVIMFLW RPSA 106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFFGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP 107 1457 A 1386 719 558 FFFVTRSISVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFGGFGAAVGCCV/GGFLGDWLGTRK							t I
103 1453 A 1371 2 410 CHSTESSDFILPGDYLLGGLCPLHSGCLQV/C SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ  104 1454 A 1376 3 432 NSRVEDRS/MSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGPQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTRIAK CQSDWMERWVDDAFWSFLFSLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFFGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							1
103							VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW
103			. '				
SFNEHGYHLFQAMRLAVEEINNSTALLPNITL GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLLEQ  104 1454 A 1376 3 432 NSRVEDRS/NISLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAHHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLKKNTVKF SLYRHKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFFGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK			L				
GYQLYDVCSDSANVYATLRVLSLPGQHHIEL QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP FLVPMLEQ  104 1454 A 1376 3 432 NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCL.CRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEBGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	103	1453	A	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
QGDLLHYSPTVLAVIGPDSTNRAATTÄALLSP   FLVPMLLEQ							SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
104 1454 A 1376 3 432 NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWOPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFFGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPFFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK						ļ .	
104 1454 A 1376 3 432 NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	1						
FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IIILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWTIKTFRLAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP 107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	104	1454	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD III.AVIDSIFVWFIFISLAQTIMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
KSREERFCNENTPCPVPIF  105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
105 1455 A 1379 2 396 GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD III.AVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEBGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
III.AVIDSIFVWFIFISLAQTMKTLRLRKNTVKF SLYRHFKNTLIFAVLASIVFMGWITKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1	105	1455	A	1379	2	396	
CQSDWMERWVDDAFWSFLF\SLILIVIMFLW RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP 107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK					İ		
RPSA  106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
106 1456 A 1383 1 432 EDGHGGWSSRCLVDHAEEGHREPWKRLCIW QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPFFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP 107 1457 A 1386 719 558 FFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK			L.,				
CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	106	1456	Α	1383	1	432	
FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF QNANPSNAAHSEDQPTP  107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
QNANTSNAAHSEDQPTP   107   1457   A   1386   719   558   FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL   108   1458   A   1397   631   2   RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	] i						
107 1457 A 1386 719 558 FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN YPALSLQSSWDHRHTWLIFAFL 108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							• •
YPALSLQSSWDHRHTWLIFAFL  108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
108 1458 A 1397 631 2 RVAISLLCAAIFISFMVQSAGKRWPTGVMLM VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	107	1457	A	1386	719	558	
VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK							
VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK	108	1458	Α	1397	631	2	
		!					
AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF							
	L						AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

NO: of policy   No: of polic	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Duck-took   Pepide   Column		NO: of	hod		beginning	nucleotide	
Seq	nucl-	peptide	l	in		location	
1460   1470   1485   1485   1486	1	1					
1462   1462   15   387   1464   1465   1465   1466   146		, .	l				
		"""					
	ualte	1	1	714			
Peptide							
	ĺ	ĺ	İ	i		sednence	
109	l		1				
1459				ŧ	sequence	ļ	
1459							FQQMLGQGIAGILPKLIGGYFDTDQRAAGLG
199	1			ļ	ì	•	FTYNVGALGGALAPIIGALIAORLDLGTALAS
1459	ì	ì	l	ì	1	ł	
WSHNNSNACWGRDCPYSGCKEALRTDGE	109	1459	A	1402	15	387	
RYTSRKSAKYRLQGTIPRGDVSLTILINPESSOR	'''	١٠٠٠/		1	1.5	1 301	
110	]				•		
110		1	1	i .	ĺ		CANCODE AND AND AND AND AND AND AND AND AND AND
LAELAFPYGVLATCA'SLLSCYVCVILPCSCF   FHSPDALFSLLLSCYTPSYCFYYLFSSSPL   CLLLASSPFPLATSLL LSCYTPSYCFYYLFSSSPL   CLLLASSPFPLATSL   CLLLASSPFPL		1	<del> </del>				
	110	1460	A	1421	3	350	
CLLLASSPPLEILASI   111	ľ			1		}	LAELAFPVGVLATCA*SLLSC*YCVILFPCSCF
111	1			1			FFHSPDALFSLLLLSCYFPSYCFFYYLFFSSSPL
	1		i	i			CLLLASSPFPLFILLASL
	111	1461	A	1426	2	344	FTSTMTKPFEKESEOPA*ATLAFGAOTSTTAD
TSSSNPPVAITVEGGSDEVSSYGFVNTAESST   SDELFSQOSDKLATTS   SDELFSQOSDKLATTS   STESSED   S	· '						
112	i .		1				
112			ł	1			
SRRSSQRTCSRACSGAWSRTW*RSS*TSSSSC   STSCSSSSSRSCQRPGGPLGARGVHITSCLNSC   STSCSSSSSRSCQRPGGPLGARGVHITSCLNSC   MSSSTTSSTSTT   MSSSTTSSTSTT   MSSSTTSSTSTT   MYLGENRULIDFTKGFLRGQISEMLKTR   GIFLTLSNFLIVCVLLFYVSFYLFQSCNFVL   MYLGENRULIDFTKGFLRGGISEMLKTR   GIFLTFLSNFLIVCVLLFYVSFYLFQSCNFVL   GIFLTFLSNFLIVCVLLFYVSFYLFQSCNFVL   GIFLTFLSNFLIVCVLLFYVSFYLFQSCNFVL   GRTKVHLPGHKTGPAVAKDTPEPVKKEFTVP   ATSOGP*SFFSEEPPLPPSNEVPPTLPP*EPQS   EDP*KNA*LKQMHAATTHWQQHQQHQVGC   QYHGIMQ   GRKCGQGVWPLEKDSRIFGGFLTVSNLGVEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   NYCN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MNHYKKSTLEILNPEVNPGFFLTWKQGEN   MSSGGPKVLVTERPFWQNPLTPVNGQQAG   MSSGGPKVLVTERPFWQNPLTPVNGQQAG   MSSGGPKVLVTERPFWQNPLTPVNGQQAG   MSSGGPKVLVTERPFWQNPLTPVNGQQAG   MSSGGPKVLVTERPFWQNPLTPVNGQAG   MSSGGPKVLVTERPFWQNPLTPVNGQAG   MSSGGPKVLVTREPFPCDQPL   MPWIEGHTCLDNNIHQAASEPINNFAESKR   MLAFLATGVVRHMKLEMGANLEGFGFTVS   MSLAFLATGVVRHMKLEMGANLEGFFTVS   MSLAFLATGVVRHMKLEMGANLEGFGFTVS   MSLAFLATGVVRHMKLEMGANLEGFGFTVS   MGESPAV*GYVVLAGMNSAGLSFGGGAGKY   LAEWWYNGLPGAKQALELLGKWK   MMSSFTMYLAGAGARWMEKHGFGRPKYFVP   MDLDVVGHIVHTGMLNEGGGYHCGLK   KSPVVDDMSSFTEFEITSTGDQALEVLQYLFS   NDLDVVGHIVHTGMLNEGGGYHCGLK   KSPVVDDMSSFTEFEITSTGDQALEVLQYLFS   NDLDVVGHIVHTGMLNEGGGYHCCSKAL   NKRSFFMSPFTDQQVHCWAWLKKHMPKDSN   MSCSPTMYDFTQMAWKKHGFGRPKYFVS   MDLDVVGHIVHTGMLNEGGGYHCDSLAH   NKRSFFMSPFTDQQVHCWAWLKKHMPKDSN   MSCSPTMYDFTQMAWLKKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMYDFTQMAWLKHMPKDSN   MSCSPTMAMA MAKKHMPKDSN   MSCSPTMAMA MAKKHMPKDSN   MSCSPTMAMA MA	112	1467		1424	42	272	
113	112	1462	A	1434	46	3/2	
MSSSTTSSTTSTF			ľ				
113	1			<b>!</b>			STSCSSSSSRSCGRPGGPLGARGVHITSCLNSC
MYLGEIVRNILIDFTKKGFLLRQDISEMLKTR   GIFLITFLLSNFLVCVLLFYVSFYLFQSCINVTL	<u> </u>		i	i			MSSSTTSSTTSTF
114	113	1463	Α	1439	3	292	HEDIMTHYDRLVDE*ALNAGKQRYEKMISG
114							MYLGEIVRNILIDFTKKGFLLRGOISEMLKTR
1464							GIFLTFLLSNFLIVCVLLFYVSFYLFOSCINFVI.
	114	1464	A	1463	1	396	KOOAVPEPHSSTTTPOEOEONWYGODI I NII O
115		* * * * *	<b>'</b> '	1405	•	3,0	
EDP*KNA*LKQMHAATTHWQQHQQHQVGC QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGIMQ   QYHGKSCRQYWPLEKDSRIRFGFLTVSNLGVEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN   MNHYKKSTLEILNPEVNPGFFLTLWKQGEN   MNHYKYQVWEWLQHFLDTNQLDANCIPFQEF   DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ   HLKWNGDSLFLCLSLPC   DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ   HLKWNGDSLFLCLSLPC   QUASTSYPPVCMPLNNTQKSKQPLPSAPEN   NPEEFLASDPNNEESL*RPWALEDFEIGRPLG   KGK   GYSSNSFQVPYVCMPLNNTQKSKQPLPSAPEN   NPEEFLASDPNNEESL*RPWALEDFEIGRPLG   KGK   GYSSNSFQVPYVCMPLNNTQKSKQPLPSAPEN   NPEEFLASDPNNEESL*RPWALEDFEIGRPLG   KGK   GYSSNSFQVPYVCMPLNNTQKSKQPLPSAPEN   NPEEFLASDPNNEESL*RPWALEDFEIGRPLG   KGK   GYSSNSFQVPYVCMPLNNTQKSKQPLPSAPEN   NPEEFLASDPNNFAESKR   NLAFLATGVVRHMKLFMGANLEGPGPTVS   H   H   H   H   H   H   H   H   H		)					
QYHGIMQ				i .			
115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYSFEEG GRRKCCQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN  116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPFFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KFPPCDQPL QHPWIEGHTCLIDNIHQAASEPINNNFAESKR NLAFLATGVVRHMKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGFFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFIKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHKVMEVMPHYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKHMPKDSN							
GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN  116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHNIHLQ YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSIFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIINSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSTFIEFITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							
MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN NYCN	115	1465	A	1464	291	2	
NYCN							GRRKCGQYWPLEKDSRIRFGFLTVSNLGVEN
116							MNHYKKSTLEILNPEVNPGFFFLTLWKQGEN
YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSFLCLSLPC  117 1467 A 1479 1 381 GTSGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							NYCN
YWTKYQVWEWLQHFLDTNQLDANCIPFQEF DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYTVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	116	1466	A	1465	667	337	LPPORPA*TDSYSTCNVSSGFLAGOSHNIHI.O
DINGEHLCSMSLQEFTRAAGTAGQLLYSNLQ HLKWNGDSLFLCLSLPC  117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN					1.41		
HLKWNGDSLFLCLSLPC	1						
117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQAQR VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLEILSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVUDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							
VLGPSNSFQRVPLQAQKLVSSHKPGQNQKHK QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVWWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	117	1467		1470	1	201	
QLQATSVPHPVCMPLNNTQKSKQPLPSAPEN NPEEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	'''	140/	A	14/3	4	201	
NPEELASDPNNEESL*RPWALEDFEIGRPLG KGK  118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							
1468	] ]						
118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							NPEEELASDPNNEESL*RPWALEDFEIGRPLG
PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN					<u></u>		KGK
PYWDDMSDSAKHFIRPLTGRDP*KPFPCDQPL QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	118	1468	Α	1485	3	385	TYLWL*GNPPFYEKNDGGLFELILRAKDEFNS
QHPWIEGHTCLDNNIHQAASEPINNNFAESKR NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							
NLAFLATGVVRHMRKLFMGANLEGPGPTVS H  119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN						İ	
H   1469   A   1486   I   398   GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL   NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE   KALTKFLKWVNWDLPQEAKQALELLGKWK   PMDVKDSLELLSSHYTNPTVRRYAVARLRQA   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLMYL   DDEDLLWYL   DDEDL							NI AFI ATGUVRHMDEI EMGANI EGEGETTIC
119 1469 A 1486 I 398 GTTSKHH*LARSLIRGPFDHDLKPNAATRDQL NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	( l	- Y					
NIIVSYPPTKQLTYEEQDLGWKFRYYLTNQE KALTKFLKWVNWDLPQEAKQALELLGKWK PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	110	1460	_	1496	<del></del>	200	
KALTKFLKWVNWDLPQEAKQALELLGKWK   PMDVKDSLELLSSHYTNPTVRRYAVARLRQA   DDEDLLMYL	119	1409	A	1486	l .	398	
PMDVKDSLELLSSHYTNPTVRRYAVARLRQA DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	j [					· · ·	
DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN						ĺ	
DDEDLLMYL  120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN							PMDVKDSLELLSSHYTNPTVRRYAVARLRQA
120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGAGKY LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	L						DDEDLLMYL
LAEWMVHGYPSENVWELDLKRFGALQSSRT FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	120	1470	A	1497	3	999	
FLRHRVMEVMPLMYDLKVPHWDFQTGRQL RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN					·		
RTSPLYDRLDAQGARWMEKHGFERPKYFVP PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	(	(			i	ľ	
PDKDLLALEQSKTFYKPDWFDIVESEVKCCK EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	. <b> </b>				ļ	ļ	
EAVCVIDMSSFTEFEITSTGDQALEVLQYLFS NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	ļ ļ	.]			ļ	-	
NDLDVPVGHIVHTGMLNEGGGYENDCSIARL NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	) 1	Ì				. ]	
NKRSFFMISPTDQQVHCWAWLKKHMPKDSN	[	. [			{		
					ļ		
	_ }	1				ļ	
							LLLEDVTWKYTALNLIGPRAVDVLSELSYAP

No: of nucleoted could be sequence   Despir	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Desirion   Sequence   USSN   Ocation   Option	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Seq.   uence			<u>[</u>				I=Icoleucine K=I veine I =I eucine
121   1471   A   1498   3   306   MTPDDDDRNTYGVALVWKFQTQSLTLSDLHRXSHLWRGIVSTLLEGRIGUTORY		•					
amino acid   reidue of   peptide   sequence   T=Threonine, V=Valine, W=Tryptophan, Y=Tryptopina, Y=Unknown, **Stipp codon,   possible mulectotide deletion, V=possible nucleotide insertion   MTPDHFPSLFCKEMSVGYANGIRVMSMTHT   GEPGFMLYTPEYRWGFIMLSTLVSNS   GEPGFMLYTPEYRWGFIMLSTLVSNS   MTPDHFPSLFCKEMSVGYANGIRVMSMTHT   GEPGFMLYTPEYRWGFIMLSTLVSNS   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTICDQN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTILQTVSNSPDLSSPFVEN   WPNSPDVLNHGGFYWGCLSKDCTIGYVSNSE   MLVAHTHITVEEHGRTILQTVSNSPDLSSPFVEN   WPNSPDVLNHGGFYWGCLSCONG   WPNSPGSGTPGGGAX*PPAVEQAE*   RGHVEMMEKLTFLNLHTSEKDKGGNTALHLA   AKHGHSPAVQVLLAGWQDIDKMNEKQQTPL   HVAADRG   WARAGHSPAVQVLLAGWQDIDKMNEKQQTPL   HVAADRG   WARAGHSPAVQVLLAGWQDIDKMNEKQQTPL   HVAADRG   WARAGHSPAVQVLLAGWQDIDKMNEKQQTPL   HVAADRG   WARAGHSPAVGVLSNSHALBOLGSPGCV   WARAGHSPAVGVSKSLMPKTURQWREGOFPH   HVERRGGVDITAWDKDAGKRDDFIGRQV DLSALSREQTHKLEQLEEGGGHLVLLVTTT   ASATVSISDLSVNSLEDQKEREBELRKYSPRIFTHKLEQUEEGGGHLVLLVTTT	1 - 1			914	•		
		'	ĺ		amino acid	of peptide	
MPDIFFSLICKEMSVGYANGIRVMSMTHT   GERGFM NYPIEYRWGFMISTLSUNSS	1 1						
121				<u> </u>	sequence		
121	1						
122	121	1471		1400		206	
MLVAHTHIVEEHTGTHLQYVSWPDHSVPDD	121	14/1	А	1470	,	300	
122	1 1						
122	1						
WPPGTDTALGWSKQPSQSYTLFES*VGSGYII   DNFFLA	122	1472	A	1533	121	329	
DNFFLA	1						
RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA	L						
AKHGISPAVQVLLAQWQDINEMNEKQQTPL	123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE
124	1						
1474							
HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP   YVKFRLGHQKYKSKIMPKTILNFQWREQFDF    HILYEERGGYUDITAWDKDAGKRDDFIGRCQV    DLSALSREQTHKLELQLEEGEGHLVLLVTLT    ASATVSISDLSVNSLEDQKEREEILKRYSPLRI    FHNIKDVGFLQVKVIRAEGLMAADVTGKSD    PFCVVELNNDRLLTHTYYKNLNPEWNKVFIL    *VALVWKKFQTQSLRLSDLHRKSHLWRGIVS    ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY    KSKIMPKTI.NPQWREQFDFHLYEERGGVIDIT    AWDKDAGKRDDFIGRCQVDLSALSREQTHK    LELQLEEGEGHLVLLVTLTASATVSISDLSVN    SLEDQKEREEILKRYSPLRIFHNILKDVGFLQV    KVIRAEGLMAADVTGKSDPFCVVELNNDRLL    THTVYKNLNPEWNKVFIL    125		1454		1.55			
YVKFRLGHQKYKSKIMPKTLNPQWREQFDF	124	14/4	A	1555	l l	/45	• •
HLYEERGGVIDITA WDKDAGKRDDFIGRCQV DLSALSREQTHKLELQLEEGEGHLVLLVTLT ASALTSISDLSVNSLEDQKEREBLKRYSPLRI FHNLKDVGFLQVKVIRAEGLMAADVTGKSD PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL *VALVWKKFQTQSLRLSDLHRKSHLWRGIVS ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREBILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASA KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLSLSPDMRTFVSGACDASA KLWDIRDGMCRQSFTGHVSDINAVS LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFIGHSGDVMSLVSTPLRYFTIVTQRRGLMA WGNMFDYPQLSFELICSTIFMPSFMMGVE WPDSFM TQAINR  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCLISIDRYVGVSYPLRYPTIVTQRRGLMA	1 1						
DISALSREQTHKLELQLEEGEGHLVLLVTLT   ASATVSISDLSVNSLEDQKEREEILKRYSPLRI   FINILKDVGFLQVK VIRAEGLMAADVTGKSD   PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL   *VVALVWKKFQTQSLRLSDLHRKSHLWRGIVS   ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY   KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT   AWDKDAGKRDDFIGRCQVDLSALSREQTHK   LELQLEEGEGHLVLLVTLTASATVSISDLSVN   SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV   KVIRAEGLMAADVTGKSDPFCVVELNNDRLL   THTVYKNLNPEWNKVFTL   THTVYKNLNPEWNKVFTL   THTVYKNLNPEWNKVFTL   CGGLDNICSTYNLKTREGNVRVSRELPGHTIGY   LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT   TTFTGHSGDVMSLSLSPDMRTFVSGACDASS   KLWDIRDGMCRQSFTGHVSDINAVS   KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL   EMLPTCDLADQHNIKFHYAFALNR*ER   1477   A   1612   1   497   TESPLLVRYYLPYTTKSELHAIMTAGFSTTAGS   VLGAYISFGVPSSHLLTASVMSAPASLAAAKL   FWPETEKPKITLKNAMKMESGDSGMLL*AAT   QGASSSISLVANIAVNLIAFLALLSFMNSALA   WVGNMFDYPQLSFELICSYIFMPFSFMMGVE   WPDSFM   WPDSFM   WPDSFM   TQAINR   TQAINR   TQAINR   TQAINR   TQAINR   TQAINR   TQAINR   TQAINR   TGAINF TOT TOT TO THE TOT TO THE TOT TO THE TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TOT TO THE TOT TOT TO THE TOT TOT TO THE TOT TOT TOT TOT TOT TOT TOT TOT TOT TO	l 1						
ASATVSISDLSVNSLEDQKEREEILKRYSPLRI	]						
FHNLKDVGFLQVKVIRAEGLMAADVTGKSD   PFCVVELNNDRLLITHTVYKNLNPEWNKVFTI   *VALVWKKFQTQSLRLSDLHRKSHLWGIVS   ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY   KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT   AWDKDAGKRDDFIGRCQVDLSALSREQTHK   LELQLEEGEGHLVLLVTLTASATVSISDLSVN   SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV   KVIRAEGLMAADVTGKSDPFCVVELNNDRLL   THTVYKNLNPEWNKVFTL							
PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL							FHNLKDVGFLOVKVIRAEGLMAADVTGKSD
*VALVWKKFQTQSLRLSDLHRKSHLWRGIVS   ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY   KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT   AWDKDAGKRDDFIGRCQVDLSALSREQTHK   LELQLEEGEGHLVLLVTLTASATVSISDLSVN   SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV   KVIRAEGLMAADVTGKSDPFCVVELNNDRLL   THTVYKNLNPEWNKVFTL     125							PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL
KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFINILKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHILLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNIL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WYGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 1 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA							
AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TIFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTLAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	1						
LELQLEEGEGHLVLL VTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHILLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA							KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT
SLEDQKEREEILKRYSPLRIFHNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHILLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	j	[					
KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL  125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	i i		i				
125 1475 A 1556 57 509 GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA		j	,	J			
125	}						
CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	125	1475	Ā	1556	57	509	
LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS  126 1476 A 1592 3 178 KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER  127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA						, , ,	
KLWDIRDGMCRQSFTGHVSDINAVS		}					
126	[	ĺ			. 1	•	
EMLPTCDLADQHNIKFHYAFALNR*ER     127							
127 1477 A 1612 1 497 TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHILTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	126	1476	A	1592	3	178	
VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	107	1455		1610			
FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	127	14//	A	1612	1	497	
QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA			1				
WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM  128 1478 A 1619 286 486 CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	l	l	- 1	·	- 1		
WPDSFM			- 1				
128   1478   A   1619   286   486   CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR   129   1479   A   1627   I   395   PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA		}	l	Ì	-		
EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	128	1478	A	1619	286	486	
TQAINR  129 1479 A 1627 I 395 PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA	-		ļ				
MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA							TQAINR
	129	1479	A	1627	1	395	
LLCVWALSLVIYIGPLLGWRHPAPEDETICOI	1		1	l			
	I			l			
NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV		{	1	}		ļ	
	130	1400		1620		166	
	130	1400	^	1036	-	400	
ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS EMHSFIOIOKNTNORSHDSRSMALPOEOSOHP		1	j	İ			
KPSEASTTLPESHLKTPOMPPTTPSSSSFTKVT	, <b> </b>	·	1	]	j		
KDKDIK*LLFNLYSSVEILPEVLHLKT				1		ļ	KDKDIK*LLFNLYSSVEILPEVI HI KT
131 1481 A 1651 607 3 LAEGGDVFDCVLNGGPLPESRAKALFROMVE	131.	1481	A	1651	607	3	
AIRYCHGCGVAHRDLKCENALLQGFNLKLTD	. 1	Į	1	· (			
FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ		İ	ŀ				
GIPHDSKKGDVWSMGVVLYVMLCASLPFDD	1		1		1		GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TDIPKMLWQQQKGVSFPTHLSISADCQDLLK RLLEPDMILRPSIEEVSWHPWLAST**KOWOV
132	1482	A	1656	150	48	LSNKVGGESKPKKKK  LVAKSLLYCGCLFFLLQLAKNVGNNSFNDIM EANLTSPSPKPTPSSDM*VFLIY*TYFGAWHV VDAQ
133	1483	A	1660	3	406	RKHIKLLIQKLSDVP*ECQNNQL*KLTEICEKE KKEFKKKMDDQRPEKITEA*SKDKSPMEEEK TEMIRSYIQEVGRYIKRLEEAQSKRLEKLREK HKEIRQPILDEKPKGEGSSSFLSETCHEDTSWF PNFTP
134	1484	A	1666	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPP FFPAGAPPASSSSSSSSSSSPPTVSTAPPLIPPPGF PPPPGAPPPSLIPTIESGHSSGYDSRSARAFPYG NVAFPHLPGSAPSWPSLVDTSKQWDYYARSS SSSSSSSSSSSSSPRDRDRER*RTREREREDHS PTPSVFNSDEERYRYREYAERGYERHRASRE KEERHRERRHREKEETRHKSSRSNSRRRHESE EGDSHRRHKHKKKSKRSKEGKEAGSEPAPEQE STEATPAE
135	1485	A	1673	i	417	PTRPVNSSQAFALVYYTLGALGGNLIAHMGL GYRYWAGIGVLQSCESALTHYRLVANHVAS DISLTGGSVVQRIRLPDEVENPGMNSGMLQE DLIQYYQFLAEKGDVQAQVGLGQLHLHGGR GV*QNHQRAFDYFNLAA
136	1486	A	1678	525	9	ANTSLSSAAVSAVSPPPCRTSTATTLPPPMPSF FCVFPSPSMSPSPSEFLSCIASVSRVHSLSSSSS GSSSTASSLNFSAIMGSSSATASWVLSTASTPP CPSALPSSPAQES*SLAASSSAWPVAGISPSGA CTFPAGSASGAAKAPSPSWRCPSFRALFSLLD SSSLSL
	1487	A	1680			AHRDEIQRKFDALRNSCTVITDLEEQLNQLTE DNAELNNQNFYLSKQLDEASGANDEIVQLRS EVDHLRREITEREMQLTSQKQTMEALKTTCT MLEEQVMDLEALNDELLEKERQWEAWRSVL GDEKSQFECRVRELQRMLDTEKQSRARADQ RITESRQVVELAVKEHKAEILALQQALKEQK LKAESLSDKLNDLEKKHAMLEMNARSLQQK LETERELKQRLLEEQAKLQQMDLQKNHIFR LTQGLQEALDRADLLKTERSDLEYQLENIQV LYSHEKVKMEGTISQQTKLIDFLQAKMDQPA KKKKVPLQYNELKLALEKEKARCAELEEALQ KTRIELRSAREEAAHRKATDHPHPSTPATARQ QIAMSAIVRSPEHQPSAMSLLAPPSSRRKESST PEEFSRRLKERMHHNIPHRFNVGLNMRATKC AVCLDTVHFGRQASKCLECQVMCHPKCSTC LPATCGLPAEYVTHFTEAFCRDKMNSPGLQT KEPSSSLHLEGWMKVPRNNKRGQQGWDRK YTVLEGSKVLIYDNEAREAGQRPVEEFELCLP DGDVSIHGAVGASELANTAKADVPYILKMES HPHTTCWPGRTLYLLAPSFPDKQRWVTALES VVAGGRVSREKAEADAKLLGNSLLKLEGDD RLDMNCTLPFSDQVVLVGTEEGLYALNVLK NSLTHVPGIGAVFQIYIIKDLEKLLMIAGEERA LCLVDVKKVKQSLAQSHLPAQPDISPNIFEAV KGCHLFGAGKIENGLCICAAMPSKVVILRYN ENLSKYCIRKEIETSEPCSCIHFTNYSILIGTNK FYEIDMKQYTLEEFLDKNDHSLAPAVFAASS NSFPVSIVQVNSAGQREEYLLCFHEFGVFVDS

NO. of   No. of   N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Deptide	NO: of		1				D=Aspartic Acid, E=Glutamic Acid.
Sequence			1				
149   1499   A   1814   1494   A   1814   1496   A   1827   26   448   XVEEKIADURES ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR ASSOCIATION   XVESTOR   XVES	eotide		ł				
148		1 •	Į.				
mino acid residue of peptide residue of peptide residue of peptide sequence   T-Threonine, V-Valine, W-Typtophan, Y-Tyrosine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova, X-Siop codon, peptide sequence   T-Threonine, X-Valnova,			J				
				7.1			
peptids	}	ł		1			
						Sequence	
148		1	1				
NSLEVIEIQARSSAGTPARAYLIDIPNEYLICPA	<del></del>	<u> </u>	├		sequence	ļ	
ISSGATYLASSYQDKLRVICKGMLVKESGTF   HIRGESTSRRP-PASPLPQYGGGRALGGSLV   HIRGESTSRRP-PASPLPQYGGGRALGGREK   HIRGESTSRRP-PASPLPQYGGGRALGGSLV   CLCGGAAGRI-ASSGKSLGRRPEGGGAC   CWPRICCIGGAAGRI-ASSGKSLGRRPEGGAC   CWPRICCIGGAAGRI-ASSGKSLGRRPEGGAC   CWPRICCIGGAAGRI-ASSGKSLGRRPEGGAC   CWPRICCIGGAAGRI-ASSGKSLGRRPEGGAC   CWPRICCIGGAAGRI-ASSGKSLGRRPEGGAC   PONGSCASASQBAAPPPMCPPGRRWAVAS   PETECAAPGTICRILEAA   1890			l				
HHRGPSTSRR*PÅSPLPOYGGGRÄLGGSLVW	1	ł	İ	!			
1488		!	l	[		1	
PLCCLGGAAGRI-ARSGKSGLRRRAHAGDE   PLCCLGGAAGRI-ARSGKSGLRRRAHAGDE   PGCPONSCP'CSAPSGGGGUPLEPGTGGVCS   CWTRGCQTTARTAAAAAAPGPAGRRPPGGA   PQNGSCAASASGEAAAPPPMCPFGRWAVAS   PETRECPAAPGTRGRRLEAA     139							
	138	1488	A	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
1489	1		· ·				PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
1489			1				PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
1489	( I		1				CWTRGCOTTARTAAAAAAPGPAGRRPPGGA
1489			Ì				
1489	<b>!</b>		ì			ł	
140	139	1489	A	1693	3	376	
140	1			10,5	_	3.0	
140							
140	'						
HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   KNLEQREIMVS*BGCKLISQLSRGKINTWK   MAAL*FELSGRLRSSKSNGWNGDNSTGYLTV   PLRPLTVKEVYTMDVPAPNVRCINWMG   KNLEQREYTMDVPAPNVRCINWMG   KNLEQREYTMDVPAPNVRCINWMG   KNLEQREYTMDVPAPNVRCINWMG   GLQLQGGTATETLSSPTIMGRRGRREBKISS   LSLASSTVGPGGQIVHTETTEVVLCGPPLSGF   GLQLQGGTATETLSSPTIMGRRGRREBKISS   LSLASSTVGPGGQIVHTETTEVVLCGPPLSGF   GLQLQGGTATETLSSPTIMGRRGRREBKISS   LSLASSTVGPGGQIVHTETTEVVLCGPPLSGF   GLQLQGGTATETLSSPTIMGRRGRREBKISS   LSLASSTVGPGGQIVHTETTEVVLCGPPLSGF   GLQLQGGTATETLSSPTIMGRRGREBKISS   LSLASSTVGPGGQIVHTETTEVVLCGPPLSGF   GLQLQGGTATETLSSPTIMGRRGRGRAPAQCILPDA   ALAKVV   A	140	1400		1704		276	T W A GIA KO VO I Q KII G K V HL G E Q K A L G L
141	140	1490	A	1/04	3	3/6	
LVLVEVVKHLSLGTVVHCNGKMRPPEP							
141	}					}	
DKLELELVLKGSYEDTQTSFLGTASAFRFHY   MAAL*TELSGRLRSSKSNGWGDNSTGYLTV   PLRPLTIVKEVTMDWAFMVNGUNWMG   1402   1492   A 1769   1   406   NNPSTLPRGS*PMSPRTTMGRRQRREEKSS   LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF   GLQLQGGFATETLSSPFLVCFEPDSPAERCG   LLQVGDFVLSINGIATEDGTMEEANQLLRDA   ALAHKVV   ALAHKV   ALAHKVV   ALAHKVV   ALAHKV   ALAHKVV   ALAHKV   ALAKKV   ALAKKV   ALAKV   AL							
MAAL*FELSGRLRSSKSIGWNGDNSTGYLTV   PLRPLTIVKEVTMDVPAPNVRGLNWMG     142	141	1491	Α	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
MAAL*FELSGRLRSSKSIGWNGDNSTGYLTV   PLRPLTIVKEVTMDVPAPNVRGLNWMG     142					,		DKLELELVLKGSYEDTOTSFLGTASAFRFHY
PLRPLTIVKEVTMDVPAPNVRGLNWMG	1						
142							
LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF   GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG   LLQVGBRVLSINGIATEDGTMEEANQLLRDA   ALAHKVV   ALAHKVV   QMLRNGGQDTVPDYHFADRIRELL*PTEDQ   KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA   NELCEVNRKGCTISGPPCLPYFCVGGCKL.GQA   SPFIARQGTLQVPSSAGEVECYKICSCGQSGL   LENCMEMHCMDLPTDTSALVR   LENCMEMHCMDLPTDTSALVR   PGRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL   PYFLQEPQDATVNKKPVELRCRAFPATQIFF   KCNGEWVSQNDHVTQEOLDEATGLRVREVH   IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK   SRRAYVRI   LEVSRQQVEELFGLEDYWCQCVAWSSAGTTK   SRRAYVRI   LEVSRQQVEELFGCEGSE   THILALCPIVQHPEDTCHSSREVGVCCSXTIDV   RLVNGKSQCDGQVEINVLGHWGSLCDTHWD   PEDARVLCVLILTWPSWTYNTHILLISQCH   VRLYKLTHAEHDHLASIMNKLLTNYDNLFE   TSVTYSMG*HGAPTGSAGANWKH**LIAH   YYYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE   Q	142	1492	A	1769	1	406	
143	1			1,0,	·	, ,,,,	
LLQVGRVLSINGIATEDGTMEEANQLLRDA		i			;		
ALAHKVV	1						
143						•	
KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA   NELCEVNRKGCTSGDPCLPYTCVQGCKLQQA   SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL   LENCMEMHCMDLPTDTSALVR   LENCMEMHCMDLPTDTSALVR   PGRRFRPLSQAGTDSGS*VFPDSFPSAPAEPL   PYFLQEPQDAYTIVNKPVELRCRAFPATQIYF   KCNGEWVSQNDHVTQEGLDEATGLRVREVH   LEVSRQQVEELFGLEDYWCQCVAWSSAGTTK   SRRAYVRI   SRRA	1/12	1402	_	1790		447	
NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA   SDFIARQGTILIQVPSSAGEVECYKICSCGQSGL   LENCMEMHCMDLPTDTSALVR     1494	143	1423	^	1707	1	447	
SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL   LENCMEMHCMDLPTDTSALVR							
LENCMEMHCMDLPTDTSALVR	1						
1494	i :		· .				
PYFLQEPQDAYIVKNKPVELRCRAFFATQIYF KCNGEWVSQNDHVTQEGLDEATGLRVREVH IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK SRRAYVRI  145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTILFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQELLGCP							
KCNGEWVSQNDHVTQEGLDEATGLRVREVH   IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK   SRRAYVRI     145	144	1494	Α	1814	1	404	
IEVSRQQVEBLFGLEDYWCQCVAWSSAGTTK SRRAYVRI  145 1495 A 1827 26 448 XVEKHADTWRSXCLSDFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQELIGCP							
SRRAYVRI  145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	1 1						
SRRAYVRI  145 1495 A 1827 26 448 XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
THLALCPIVQHPEDTCIHSREVGVVCSRYTDV RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146							
RLVNGKSQCDGQVEINVLGHWGSLCDTHWD PEDARVLCRQLNCGTAL  146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							
PEDARVLCRQLNCGTAL     1496	1 1			1	ŀ		
146 1496 A 1828 574 333 QHEGGDLRRRQLGEIQLTVRYVCLRAASAC* SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP			ĺ	·			
SMAAET*HHVPASGADPYVRVYLLPERKWA CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTIFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	146	1496	Δ-	1828	574	333	
CRKKTSVKRKTLEPLFDET  147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	.70	4770	^	1020	3/4	ددد	
147 1497 A 1855 1 372 ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	j l			]		ĺ	
VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	142	1400		1055	<del>-,</del> -		
TSVTYSMG*HGAPTGSEAGANWNH**LHAH YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	14/	149/	A	1822	1	<i>5T2</i>	
YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE Q  148	( 1		- 1	ľ			
148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP			. 1	l			TSVTYSMG*HGAPTGSEAGANWNH**LHAH
148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	ļ <b>1</b>		·	ł	j		YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
148 1498 A 1879 568 7 LLSALDDKGGTQPSASFSNAPTIVCVTACPAG IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP							Q
IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS 149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	148	1498	A	1879	568	7	
GIEGRLTADQLNSATACIFAAEVAIKESERFN GIPALSVPVAEPIRHAEALMQQALTLKRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP				1			
GIPALSVPVAEPIRHAEALMQQALTI.KRSDET RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	1 1		. [	1	Ĭ	İ	
RTVQQDTQPVKSVKTELKQALLSGISFAVPLI VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 1499 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	[		!				CIPAL CVPV AFPIRHAFAL MOOAL TI VECTOR
VAGGTQVA*AV*RQGISSLHDVQVRTWNS  149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	1. 1	ł	ì	[	!		RTVOODTOPVKEVETEI KOALI GOIGEANDEL
149 A 1880 611 24 GLNSENALSNEAMERGWQCLRLFAERLQDIP PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	j.						
PSQIRVVATATLRLAVNAGDFIAKAQEILGCP	140	1400	<del>  </del>	1000		74	
	147	1477	Α :	1000	011	24	
VQVISGEEEARLIYQGVAHTTGGADQRLVVD	} }	J	J	)	}	j	
	L					l	VQVISGEEEARLIYQGVAHTTGGADQRLVVD

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	j	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	] .	ļ	}	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	Ì	i		peptide	coquance	/=possible nucleotide deletion, \=possible
	4.	]		sequence		nucleotide insertion
ļ		<u> </u>	<del> </del>	2cd acuec	<del> </del>	IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
i I					}	YFADRNLGQENFDAAQKAAREVLRPVADEL
		l	1			RYHSWKEVRGASVTVOALOEIMMAOGMDE
i i	ł	ì	l		·	RITMEIWPVD
150	1500	A	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
130	1500	) ^	1074	2	1 730	LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
1		l			Ì	EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
	Į	1			ļ	PESSILDGMIRQLQQQQDQRMGADQDTIPRG
	i	1				LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGOV
ļ					ļ	EGVRQMHQNAPRSQIATERDLQAWKRRVVV
ļ		j			j	PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
151	1501		1900	141	785	TLQLSHKSDSVGLVSQSRPRTCRRKYP
121	1501	Α	1900	141	/63	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA
	,			İ		MEMQIKKQFQDTCKVQTKQYKALKNHQLEV
}	1	ł	ł		ł	TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
			1			NEMMASQALRLDEAQEAECQALRLQLQQEM
1		İ				ELLNAYQSKIKMQTEAQHERELQKLEQRVSL
			,			RRAHLEQKIEEELAALQKERSERIKNLLERQE
	4.500					REIETFDMESLRMGFGNLVTLDFPKEDYR
152	1502	Α	1915	2	377	LVRLLDTQRDGLQNYEALLGLTNLSGRSDKL
] .					ļ	RQKIFKERALPDIENYMFENHDQLRQAATEC
·	ĺ					MCNMVLHKEVQERFLADGNDRLKLVVLLCG
ļ			}			EDDDKVQNAAAGALAMLTAAHKKLCLKMT
			1.55			QVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL
		i			-	ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG
					l 	FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
1						YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
			[			RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
						QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA
1						KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
l l					'	LLLKDMDLLRGREAELKQRVEAFESYQLELK
L						DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	À	1935	1	270	VTRKLPIFÍVDAFTARAFRGSPAADČLLENEL
<b>}</b>			1			DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
	_					RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	Α	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
	•					AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
L						NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	A	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA
.						NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
[						RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
			}			QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
[						RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
						NGHLATVKLLVEEKADVLARGPLNQTALHL
] 1						AAAHGHSEVVEELVSADVIDLFDEQGLSALH
	·					LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
						HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
						IYSEYCNNHPGACLELANLMKQGKYRHFFEA
		-				CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL
	1					LKYTTQEHGDYSNIKAAYEAMKNVACLINER
						KRKLESIDKIA
161	1511	A	1984	4	770	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
						LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
						SSHNASKTLNANNMETLIECQSEGDIKEHPLL
						TITLE TOTAL TOTAL TOTAL CONTROL TOTA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Ghrtamine, R=Arginine, S=Serine,
		[	!	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	<b>i</b> .			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ	<b>,</b>	peptide	{ -	/-possible nucleotide deletion, \-possible
			!	sequence		nucleotide insertion
						ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
			]		i	PASDFSGALETDLKASLFDQPLSIICGDSDTLP
j			]		J	RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
					1	KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
L						RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	Α	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
1						KFQGRWGTVCDDNFNIDHASVICRQLECGSA
1			Ì			VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
						CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	Α	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPQPAATNGD
1 1						LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
						FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
						SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
	•					ENNWYFVVADSSKAGFTTIYKWERETGFYSH
					100	QSFTR
165	1515	A	2013	2	403	EDPEELGHFYDYPMALFSTFELFLTIIDGPANY
						NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
j l						DTHWRVAHERDELWRAQIVATTVMLERKLP
1 1						RCLWPRSGICGREYGLGDRWILRVEDRQDLN
166	1516	_	2019		000	RQRIQRYA
100	1510	A	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
						QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
1						NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
1 1				·		SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI WDMRNLATIFLAVVMALLSLHCLAAFKRLE
1 1						HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
] ]						VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
						LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
j [						GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
						HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA
1						ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
						PEVVLPDPVEETRHHAEVVKKVNEMIVTGOY
1 1	}		ļ .			GRLFAVVHFASRQWKVTSEDLILIGNELDLA
	ļ	- 1				CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
) ]	j	i		ļ		RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
L!						TTPQTVLRINSIEIAPCLL
168	1518	Ā	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
		ĺ	(	ĺ		RLQGAARVFMPLQAQVKAKASKPLQMQIKA
	ļ			ļ	i	PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
	1					KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	Α	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR
				ļ		TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
	ļ	1		}	ļ	EDNSRSKREGLFHENECIVKINNVDLVDKTFA
	- 1	[	l			QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
	ŀ	ŀ	ł	ļ	ļ	VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
1					}	NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
	ļ		Ì	J		SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
	j			}		TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ
, 1	1	į	ŀ	Ì	j	SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
170	1500		2056			ETASLVIARQEGHFLPRELVMFRSQSH
170	1520	A	2050	363	1	PVATHLTKILNSDEHAVVISSAKTLCETVKDF
	!	ļ	Ì	i	İ	VAKVEKTYDKTLENAVVADAVASKCSVLNE
	j	. <b>i</b>		1		KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
171	1521	<del></del> _	2055	120	675	ESSSEESLGESKEQLGDDVTKPSSQKA
1/1	1521	A .	2055	139	675	IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS
ļ ļ		ł			l	DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
<b></b>			<u>_</u>			LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

WO 01/57188 PCT/US01/03800 .

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion RESCTITAYPCDSYQDYRNGKCVSCGTSQKE
				···		SCPLLGYYADNWKDHLRGKDPPMTKAFFDT AEESPFCMYHYFVDIITWNKNVR
172	1522	A	2056	3	361	LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK MEWKVKIRSDGTRYTTKRPVRDRILKERALKI KEERSGLTTDDDTMSEMKMGRYWSKEERKQ HLVRGKEQRRRREFMMRIRLKCLKES
173	1523	A	2060		387	GTRILSMQIPFVGFQPIRTSEHMAAAGVFALL QAYAFLQYLRDRLTKQEFQTLFFLGVSLAAG AVFLSVIYLTYTGYIAPWSGRFYSLWDTGYA KIHIPIIASVSEHQPTTWVSFFFDLHILGCTFPA
174	1524	A	2071	74	443	G LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE RRLKEEEEARLKYEKEEMERLEIQRIEKEKW
175	1525		2000	139	404	HRLEAKDLERRNEELEELYLLERCFPEAEKLK QETKLLSQWKHYIQCDGSPDPSVAQEMNT
1/3	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV AESSTVGWLCALFRVTHVGVGATGHGVVCG RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ RLQFPYLEPGHELPATTLLAFLAAV
176	1526	A	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT GIHSVYTVYQGHEIMFHVSTMLPYSKENKQQ VERKRHIGNDIVTIVFQEGEESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET
177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC CDGAWLAWACWVFGNDFPSPASAACSALLG CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC
178	1528	A	2104	2	409	RSDLQ ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD
179	1529	A	2111	1	312	PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG
180	1530	<b>A</b>	2116	3	366	TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV
181	1531	A	2117	2	386	YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI
182	1532	A	2123	1	493	RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMOFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS
183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Veline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	A	2145	3	538	HELTVAAADRGQPPQSSVVPVTVLDVND NPPVFTRASYRVTVPEDTPVGAELLHVEASD ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA PVTIEVQDVNDHGPAFPLNLLSTSVAENQPPG TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A .	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW PKNNFNGSLVQASYQHEELRREVIMLACSFG NKHCHQQASTLISDWISSNRNRIPLNVRDIVY CTGVSLLDEDVWEFIWMKFHSTTAVSEKKIL LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI DVIIHVARNPHGRDLAWKFFRDKWKILNTRI RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYTVHPLWETWA HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP LDEQNRDWQGLLENLHVELTLDEEDSEGPEK EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC AHDKSP
187	1537	A	2158	227	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDQPG OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP AVVVPYMMVLQENGYGVEEGIPTLLMAASS MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV AIVVSSLDW
190	1540	A	2179	64	399	MRLNONTLLLESFGXXRPYTSEHAPTYHQW MKADELLRWTTSEPLTLEHEYAMQRTWLED AYECTFIVLDAEKRHAQPGATEESCMVGDVN LFLTDLEDLTLGEIEVLIAEP
191	1541	A	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR LSYVLFIQERDVHKGMFATNVTENVLNSSRV QEAIAEVAAELNPDGSAQQQSKAVNKVKKK AKRILQEMVATVSPAMIRLTGWVLLKLFNSF FWNIQIHKGQLEMVKAATETNLPLLFLPVHR SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS YTHSKGIMHRDVKPLNILCNSPRNKVILADW GLAEFYHPMRKYSVHVATRYYKSPEILLDYE YYDYSLDIWAVGVILLELLTLKLHVFEGGDN EQ
194	1544	A	2241	105	409	RKGVGKMPTSEGRPGQERSDWVTSYKVMGS NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE LPVPMGARYIRINPQSWFDNGSICMRMEILGC PLPDPNNY
195	1545	A	2245	1	672	MGVASDWTKRIEYQPGSGSMPLFPSIHLETCD GAVSSLQIVTELQTNYIGKGCDRETYSEKSLQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW MKHGPSPGVRAEKETILCYSDKTEMNRHHY ALYVHNCRLVFLLRKDFDQADTFRPAEFHW KLDQQALAKVDGQPGKSITRQLQEMPVTIQG ISILKPS
196	1546	A	2256	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF QPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL HCFTESGRGKSWRLCAAILPL
198	1548	A	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATQMFF FLGFASNNCFIMAAMSYDRYTAIHNPLQYHT LMTRKICLQMMMASWMVGFLFSLCIIVTVFN LSLCDLNTIQHYFCDISPVVSLACNYTFYHEM AIFVLSA
199	1549	Α	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI KRLAFCHSNVLSHSYCVHQDVMKLAYADTL PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2334	2	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH DQLQFEYKEEM
201	1551	A	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM TEELEALRSSSLGSRTLDPLWKVRRSQKLDM SARLELQSALEAEIRAKQLVQEELRKVKDAN LTLESKLKDSEAKNRELLEEMEILKKKMEEK FRADTGKLMLCDSALFEYKYFSNECFYFLFD LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL NRVLERLAGGATRDSAASDILLDDIVLTHSLF LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA. CLAMLLHFLDTYQGLLQEEGAGHIIKDLYL LIMKDESLYQGLREDTLRLHQLVETVELKIPE ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLQ PTEDCVFTALGINSHLFACTRDSYEALVPLPE EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR CVHELEFVDYVFHGE
203	1553	A	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL LKHGASLDNFHFIGGSLGAHISGFVGKIFHGQ LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ AGSRLGAMRRCAREMDATPMPPAPSCPSERV T
205	1555	Ā	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN HGQDRRPHAPGRPARGKVQEGSARPPSAVAC EDCSCR

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	[	1	[	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]	ļ	j	<b>!</b>	peptide		/=possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH
j						THIPCPVSSTTNDTPDQIWVSVGSLRMGTGG
					[	MGANASTSPRCWDLSSGNKKWIIQVPILASIV
	ļ					ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	A	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT
		L			Í	HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL
		ŀ				MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT
		ì				GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA
ì	İ				1	TLPLTLIVILENIAVAWIYGTKKFMQELTEML
		[		L		GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP
i		l			1	NASNLDKVLTDIKADKDQANDGLSSALLILY
1 1		1			ľ	LDSARNLPIRYKTNEPVWEENFTFFIHNPKRQ
1 '						DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM
1					·	TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE
i						RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG
1						SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE
1 1		ĺ				AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA
						SDISLPIATQELRQRLRQLENGTTLGQSPLGQI
						QLTIP
210	1560	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA
1 :			٠ ا			AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW
1 1						GAVRAAESLTDIAEPASPQVHETPIDASQTQK
			1			VEPASKSRFTPELQAKVSHSRERALSTMDATP
						HHAQPQRGEG
211	1561	Α	2431	1	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP
						LMFWLHGIQLVALNYQTDDLPLHLNAAMFE
1						ANGGCGYVLKPPVLWDKNCPMYQKFSPLER
1						DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE
						VDVLGMPLDSCHFRTKPIHRNTLNPMWNEQF
1 1			Ì			LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL
						KALKRGYRHLQLRNLHNEVLEISSLFINSRRM
1						EENSSGNTMSASSMFNTEERKCLQTHRVTVH
1220	1500		0436			GVPG
212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI
1 1						YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC
						TDLDHNLDKGYLTVLGEQATPTNRLGALPKG
						RANRTRDLELTYLAERIVRLTWIPGDANNRPI
1212	1562		2445		1204	TDYDCQIEEHQ
213	1563	Α	2445	1	1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT
] [	İ			ſ	1	WLMPELHPKEQILELLVLEQFLSILPEELQIWV
				İ		QQHNPESGEESVTLLEDLEREFDDPGQQVPAS
				l		PQGPAVPWKDLTCLRASQESTDIHLQPLKTQ LKSWKPCLSPKSDCENSETATKEGISEEKSQG
				j		LPQEPSFRGISEHESNLVWKQGSATGEKLRSP
				ļ	•	
[ ]						SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT TDSIMCQKVPPEERPYRCDVCGHSFKOHSSLT
			1	ł	' l	QHQRIHTGEKPYKCNQCGKAPSLRSYLIIHQR
] [				ļ		IHSGEKAYECSECGKAFNOSSALIRHRKIHTG
		ſ	ĺ			EKACKCNECGKAFSOSSYLIIHORIHTGEKPY
1	ļ	Į				ECNECGKTFSOSSKLIRHORIHTGERPYECNE
.		Į			!	CGKAFRQSSELITHQRIHSGEKPYECSECGKA
	l	- 1				FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL
	1704	^	2701	•		STRLVSVQEDAGKSPARNRSASITNLSLDRSG
1		- 1	ł	}		SPMVPSYETSVSPQANRTYVRTETTEDERKIL
1 1						LDSVQLKDLWKKICHHSSGMEFQDHRYWLR
	-	- 1		ļ	ł	THPNCIVGKELVNWLIRNGHIATRAQAIAIGQ
L			1			TIE TICLY ON TO THE PROPERTY IN THE PROPERTY I

Decide coide sequence   USSN   09495   1	SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Diocation   Diocation   Diocation   Diagnost   Diagno			l loc	1			
uence   uence   09/496   corresponding to the first of the land on the land of the land			[				
1	I						
minto acid residue of peptide residue of peptide sequence   peptide sequence   peptide sequence   peptide sequence   peptide sequence   pertia			ļ				
Peptide		<b>,</b>					
Peptide sequence	ļ	}			residue of		
				l	peptide	•	/=possible nucleotide deletion, \=possible
1565   A   2464   3   2932   GFOVESSQDCMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRT   CSLSSQGHOPGESSYFDDDTAPASILIVNA   AAPALTNPTAPVLCTAPNNTLAQKEKVPSOMR   QPPAGVISSSTPDLTCAVSTHATVPGVRISSCTPDLTCAVSTH   STYPGVRISSCTPDLTCAVSTHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHATVPGVRISSCTPDLTCAVSHATVPGVRISPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISPDLTCAVSHATVPGVRISPD	1	Ì	·				
1565   A   2464   3   2932   GFOVESSQDCMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRCSFISGO   HORGRIGRRVCSSQDSMADVFVHLRTAWFRT   CSLSSQGHOPGESSYFDDDTAPASILIVNA   AAPALTNPTAPVLCTAPNNTLAQKEKVPSOMR   QPPAGVISSSTPDLTCAVSTHATVPGVRISSCTPDLTCAVSTH   STYPGVRISSCTPDLTCAVSTHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHATVPGVRISSCTPDLTCAVSHATVPGVRISPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISPDLTCAVSHATVPGVRISPD				<del></del>			AMVDGRWLDCVSHHDOLFRDEYALYRPLOV
HGPGRHGRRVCSSQSMADVFVHLRIAWICSLAWING   CSLSGQHGPGESVSYDDDDPASILHVNA     AAPALTNYTAPVLCTAPNITAQKEKVPSGMR     QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC     TPDLTCAVSHSTVPSVCISSCTPDLTCAVSTH     STVPGVRISSCTPDLTCAVSTHSTVPGVRISSCT     TPDLTCAVSHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTCAVSTHSTVPGVRISSCTPDLTCAVSHATVPGVRISSCTPDLTC						}	
HGFGRHGRRVCSSQJSMADVFVHLRTAWFY   CSLISGQHGGESVSYEDDIDPASILHVMA     AAPALTNYTAPVLCTAPNINTAQKEKVPSGMI     AAPALTNYTAPVLCTAPNINTAQKEKVPSGMI     GRAGVRISSRTPDLTCAVSTISTYPOVRISSC     TPDLTCAVSHSTVPSVCISSCTPDLTCAVSISH     STVPVGVRISSCTPDLTCAVSTISTYPOVRISSR     TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHA     ATVPGVRISSCTPDLTCAVSTHATVPGVRISSR     TPDLTCAVSHATVPGVRISSTPDLTCAVSTHA     ATVPGVRISSCTPDLTCAVSTHATVPGVRISSR     TPDLTCAVSHATVPGVRISSTPDLTCAVSHA     ATVPGVRISSCTPDLTCAVSTHATVPGVRISSR     TPDLTCAVSHATVPGVRISSTPDLTCAVSTHA     ATVPGVRISSCTPDLTCAVSTHATVPGVRISSTPDLTCAVSTHATVPGVRISSR     TPDLTCAVSHATVPGVRISSTPDLTCAVSTHATVPGVRISST     STVPGVRISSRTPDLTCAVSHATVPGVRISSTPDLTCAVSTHAT     STVPGVRISSRTPDLTCAVSHATVPGVRISSTPDLTCAVSTHAT     STVPGVRISSRTPDLTCAVSHATVPGVRISSTPDLTCAVSTHAT     TVPGVRISSCTPDLTCAVSHATVPGVRISSTPDLTCAVSHAT     TVPGVRISSCTPDLTCAVSHATVPGVRISSTPDLTCAVSHAT     TVPGVRISSCTPDLTCAVSHATVPGVRISST     PDLTCAVSHATVPGVRISSTPDLTCAVSHAT     TVPGVRISSCTPDLTCAVSHATVPGVRISST     PDLTCAVSHATVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHATVPGVRISST     PDLTCAVSHATVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT     PDLTCAVSHAT     PDLTCAVSHAT     TVPGVRISSTPDLTCAVSHAT	215	1565	A	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGO
CSLISGGHGESVSYEDDIPARASILLIPUA   AAPALTNTTAPVICTAPNITAQEKVYPSGMR   QRPAGVRISSRTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSCTPDLTCAVSTHSTYPGVRISSR TPDLTCAVSIHATYPGVRISSCTPDLTCAVSTHSTYPGVRISSR TPDLTCAVSHSTPDLTCAVSTHSTYPGVRISSR TPDLTCAVSHATYPGVRISSCTPDLTCAVSTHATYPGV			1				
	ł	Ì	ł			1	
TPDLTCAVSHISTVPSVCISSCTPDLTCAVSTHSTVPGVRISSC   STVPGVRISSCTPDLTCAVSHISTVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHIST   ATVPGVRISSCTPDLTCAVSHISTVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHISTVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHIST   ATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVRISSCT    TPDLTCAVSHATVPGVRISSCT					AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR		
STYPGYRISSCTPDLTCAVSTHSTYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH ATVPGYRISSCTPDLTCAVSTHSTYPGYRISSC TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH ATVPGYRISSCTPDLTCAVSTHSTYPGYRISSC TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH ATVPGYRISSCTPDLTCAVSTHATYPGYRISSC TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH ATVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH ATVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTH STYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTH TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT PDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATYPGYRISSCT TPDLTCAVSTHATYPGYRISSCTPDLTCAVSTHAT TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISSCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISCTPDLTCAVSTHATY TVPGYRISC			l .				QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
TPDLTCAVSIHATVPOWISSCTPDLTCAVSIH   ATVPGWISSCTPDLTCAVSIHSTVPOWISSC   TPDLTCAVSIHATVPOWISSCTPDLTCAVSIH   ATVPGWISSCTPDLTCAVSIHATVPOWISSC   TPDLTCAVSI	1 :			Ì		} 	TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
ATYPGYRISSCTPDLTCAVSTHSTVPGVRISSR							STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH   ATVPGVRISSCTPDLTCAVSIHATVPGVRISSR     ATVPGVRISSCTPDLTCAVSIHATVPGVRISSR     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH     ATVPGVRISSCTPDLTCAVSIHATVPGVRISSR     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH     ATVPGVRISSCTPDLTCAVSIHATVPGVRISSR     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH     ATVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH     STVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH     STVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     AVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRT     TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVP							TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
ATVPGVRISSCTPDLTCAVSTHATVPGVRISSR TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH ATVPGVRISSRTPDLTCAVSIH ATVPGVRISSRTPDLTCAVSIH ATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHATVPGVRISSRTPDLTCAVSIHS ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIHS ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHS STVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHA T	[						
TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH							
ATVPGVRISSCTPDLTCAVSIHATVPGVRISSCTPDLTCAVSHATVPGVRISSTPDLTCAVSHA	J	j					
TPDLTCAVSHATVPGVRISSRTPDLTCAVSHS							· ·
ATVPOVRISSCTPDLTCAVSTHSTVPGVRUSSC	1		!				
TPDLTCAVSHATVPGVRISSCTPDLTCAVSHAT   STVPGVCISSRTPDLTCAVSHATVPGVPHISSC   TPDLTCAVSHATVPGVRISSCTPDLTCAVSHATVPGVPHISSC   TPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVPHISSCT   TPDLTCAVSHSTPTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGRIDAVSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVRISSRTPDLTCAVSHATVPGVR			ļ .	ļ			
STVPGVRISSRTPDLTCAVSIHATVPGVHISSC TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSH STVPGVCISSRTPDLTCAVSHSTVPSVHISSCT PPLTCAVSHSTVPGVRISSRTPDLTCAVSHS TVPGVPISSCTTDLTCAVSHSTVPSVHISSCT PPLTCAVSTHTTVPGVRISSRTPDLTCAVSHS TVPGVPISSCTPDLTCAVSHSTPGVRISSRTP PPLTCAVSTHTTVPGVRISSRTPDLTCAVSHS TVPGVRISSRTPDLTCAVSHSTPGVRISSRT PPLTCAVSHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTPVGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHATVPGVRISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGULTSVGOTSG PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSTH STVPGVRISSRTPDLTCAVSHATVPGVRISSRT PPLTCAVSHSTVPGVHISSCTPDLTCAVSHS RTVSSTRFPSPRVALSHVALSGDNGGEK AVEVATVVQFTVLARAVFCNVSVACEKELDHY DVLQRAYPRVGFYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIFAREDLTMF VKRHKLESELIKSAAGGLGXSYTEPNYW  218 1568 A 2489 1 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEGONCSGALNHFRSRQPTYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP PSRLFFQEPASAVASFLNGLASV VMLCRYRT FVPASSFMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVEFCTDNCCVSLIPGEVDSVAMAPA APKIPRL VQATPARHVS MCRYLNONISSLEGTNABIQML KGVJLNONISSLEGTNABIQML KGVJLNONISSLEGTNABIQHIVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEITYLRGHLDSVNDLRTQNQVLINIS LKNANAEITYLRGHLDSVNDLRTQNQVLINIS LKNANAEITYLRGHLDSVNDLRTQNQVLINIS LKNANAEITYLRGHLDSVNDLRTQNQVLINIS		1					
TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIHS STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS TVPGVHISSCTDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRTPDLTCAVSTHA TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRTP PDLTCAVSTHATVPGVRISSRTPDLTCAVSHA TVPGVPISSCTPDLTCAVSHATVPGVRISSRT PDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSCTPDLTCAVSHATVPGVRISSCT PDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSTHAT TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSCTPDLTCAVSHA TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSTHATVPGVRISSRTPDLTCAVSHA TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVRISSRTPDLTCAVSHA TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSTHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSTHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSTHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSCTPDLTCAVSHATVPSVHISSC TPDLTCAVSTHATVPGVRISSRTPDLTCAVSHAT TVPGVRISSCTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPSCHISTPTVPVDH TVRDLTCAVSHATVPSCHISTPTVPVDH TVRDLTCAVSHATVPSCHISTPTVPVDH TVRHLESSLEVISAAAGGAGAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFSRQPTVMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRLFTQGPASAVSELIOLASLVMLCRYRT FVPASSPMYHTCVAFAWAVTLVPSLVTLFVVDH HHFGRAEMRELIQTFKGHMENSSAWVVEIQ MLKCRYDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLGJTQMLRSSLEGTNAEIGHVL KEDLEKADALTFQTTMFLKSSLENTSELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEITYLRGHLDSVNDLRTQNQVLRNS LKNANAEITYLRGHLDSVNDLRTQNQVLRNS						100	
STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGVHISSCTDLTCAVSIHS   TVPGVHISSCTIDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHITVPGVRISSRTPDLTCAVSIHS   TVPGVHISSCTIDLTCAVSIHATVPGVRISSRT   PDLTCAVSIHITVPGVRISSRTPDLTCAVSIHAT   TVPGVHISSCTPDLTCAVSIHAT   TVPGVHISSCTPDLTCAVSIHAT   TVPGVHISSCTPDLTCAVSIHAT   TVPGVHISSCTPDLTCAVSIHAT   TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSVHISSCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSIHATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHISCT   PDLTCAVSHIATVPSQHICAVS   PSTATAVSQHICAVS   PSTATAVS   PSTATAVSAVS   PDLTCAVSHIATVSQHICAVS   PSTATAVS							
PDLTCAVSHISTYPGVRISSCTPDLTCAVSTHS TVPGVHISSCTTDLTCAVSHATVPGVHISSCT PDLTCAVSTHITYPGVRISSRTPDLTCAVSHAT TVPGVRISSCTPDLTCAVSTHSTTPGVRISSRT PDLTCAVSTHITYPGVRISSRTPDLTCAVSHAT TVPGVRISSCTPDLTCAVSTHATVPGVRISSRT PDLTCAVSHATVPGVRISSCTPDLTCAVSHAT TVPGVRISSRTPDLTCAVSHATVPGVRISSRT PDLTCAVSHATVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPGVHISSCT PDLTCAVSHATVPGVHISSCTPDLTCAVSTHS STVPGVHISSRTPDLTCAVSHATVPGVHISSCT PDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPGVHISSCT PDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPGVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPGVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPGVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHAT STVPGVHISSCTPDLTCAVSHATVPGVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHAT STVPGVHISSCTPDLTCAVSTHA AVEVATVVIQFTVLRAAVFKNVSVAEGKELD LTCMTITDRADDVFPVTWSSPMPSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  460 CRRICEGPQRFESTEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPENFLPLHY DYLQFAYYRVGEYVKALECAKAYILCHPDD EDVLDNYDYYESLLDDSIDASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW VKRHKLESELIKSAAEGLGXSYTEPNYW VKRHKLESELIKSAAEGLGXSYTEPNYW VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSFAPGFPGFYGLYDKULFKHDPT SANLQLVRSSGOIQEGDLVEVVLSASATFED LQRRHALTVHSYRAP  219 1569 A 2489 3 428 SRIVLLAGAAAAEGSQODREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEHKVPGPHGKWP FSRFLFYGEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRRCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATAFAMAVTLVPSLVTLFVVDH HIFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRYDNVNSQLQVLGGHKVPGPHGKWP FSRFLFYGEPASAVASFLNGLASLEGTNAEIQRL KEDLEAADALTFQTLNFLKSSLEGTNAEIQRL KEDLEAADALTFQTLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAADALTFQLNFLKSSLEGTNAEIQRL KEDLEAANAEITVLRGHLDSVNDLRTQNQVLRNS							
TVPGVHISSCTIDLTCAVSHATVPGVHISSCT PDLTCAVSTHITVPGVRISSRTPDLTCAVSHA TVPGVRISSCTPDLTCAVSTHATVPGVRISSRT PDLTCAVSTHILTVPGVRISSRTPDLTCAVSHA TVPGVRISSCTPDLTCAVSHATVPGVRISSRT PDLTCAVSTHATVPGVRISSRTPDLTCAVSHA TVPGVRISSCTPDLTCAVSHATVPGVRISSRT PDLTCAVSHATVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHATVPGVHISSCT PDLTCAVSHATVPGVHISSCTPDLTCAVSTHS STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHATVPSVHISSC TPDLTCAVSHATVPGVHISSCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHAT AVEVATVVIGETVLESCTPDLTCAVSTH STVPGVHISSCTPDLTCAVSHAT AVEVATVVIGETVLESCTPDLTCAVSTH MQVLVCQHECVRELATRPGRLSPIENFLPLHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY MQVLVCQHECVRELATRPGRLSTEPLTHF VKRHKLESELIKSAAGCIGXSYTEPNYW VKRHKLESELIKSAAGCIGXSYTEPNYW VKRHKLESELIKSAAGCIGXSYTEPNYW  218  1568 A 2483 140 383 APSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANILQLVRSSGDIQEGLVEVVLSASATFED LQRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAALASGGQGDREPVYRDCVLQ CEEQNCSGGALNHRRSROPTYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATFAFMAVTLVPSLVTLFVVDH HIFIGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRYDNVNSQLQVLIGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAADALTFQTLNFIKSSLEGTTNAEIQRL KEDLEAANSEIQMLNASLETANTQAQLANSS LKNNAEITVLRGHLDSVNDLRTQNQVLRNS							
PDLTCAVSTHITVPGVRISSRTPDLTCAVSIHA   TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT   PDLTCAVSTHSTVPGVRISSRT   PDLTCAVSTHSTVPGVRISSRT   PDLTCAVSTHSTVPGVRISSRT   PDLTCAVSIHATVPGVRISSRT   PDLTCAVSIHATVPGVRISSRT   PDLTCAVSIHATVPGVRISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGVHISSCT   PDLTCAVSIHSTVPGULTSVSQTHSSC   TPDLTCAVSIHSTVPGULTSVSQTISSC   TPDLTCAVSIHSTVPGULTSVSQTISTG	1						
TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT PDLTCAVSIHATVPGVHISSCTPDLTCAVSIHS TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVHISSCTPDLTCAVSTHS STVPGVHISSRTPDLTCAVSTHATVPSVHISSC TPDLTCAVSHHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSTHATVPSVHISSC TPDLTCAVSHHSTVPGLITSVSQTSTG  4141 FRIKSTYRGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPVTVSSPSRMPDSTLPGS RVLARLDRDFLVHISSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCGPQRTEETYPLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASEAREDLTMF VKRIKLESELKSAAAGLGKSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALAGSGQGDREPVYRDCVLQ CEEQNCSGGALNHIFRSRQPTYMSLAGWTCRD DCKYECMWVTVGLYLOGEGBLVVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVILFVVDH HHGGRAEMRELIQTTKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEPLLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEITYVLRGHLDSVNDLRTQNQVLRNS							7
PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHAT TVPGVHISSCTPDLTCAVSTHATVPGVRISSRT PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVHISSCT PDLTCAVSIHSTVPGVHISSCTPDLTCAVSTHS STVPGVHISSRTPDLTCAVSIHSTVPGVHISSCT PDLTCAVSIHSTVPGULTSVSQTSTG  216 1566 A 2477 1 414 FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRAADVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPGREEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGETVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 MOGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTILFVVDH HHFGREAEMRELIQTIFKGHMENSSAWVVEIQ MLKCRVDNYNSQLVLGGHLGNTNADIOMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKANBEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							1
TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSTHSTVPGVHISSCTPDLTCAVSTHSTVPGVHISSCTPDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSIHATVPSVHISSCT PDLTCAVSIHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSIHATVPSVHISSCT PDLTCAVSIHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSIHATVPSVHISSC TPDLTCAVSIHATVPSVHISSC TPDLTCAVSIHATVPSVHISSC MACOVATVAIGHT PDLTCAVSIHATVPSVHISSC MACOVATVAIGHT PDLTCAVSIHATVPSVHISSC MACOVATVAIGHT PROBLEM PORT AVEVATVAIGHT PROBLEM PORT AVEVATVAIGHT PORT A	ł						
PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS TVPGVRISSRTPDLTCAVSHSTVPGVHISSCT PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSIHATVPSVHISSC TPDLTCAVSIHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSIHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSIHATVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSIHATVPGVISSC TPDLTCAVSHATVPSCONTON TOTAL PTOTAVS TOTAVS TOTAVST							
TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSTHATVPSVHISSC TPDLTCAVSTHSTVPGULTSVSQTSTG  216 1566 A 2477 I 414 FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSPSRMPDSTLPG RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAQPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP 219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPYYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRL VQATPAFMAVTLVPSLVTLFVVDH HHFGREAEMRELIQTFK GHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALITQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	1						
PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH STVPGVHISSRTPDLTCAVSHATVPSVHISSC TPDLTCAVSHSTVPGLITSVSQTSTG  216	1						,
216 1566 A 2477 1 414 FRIKSYRGYSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNYSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARSYHLL VRDVSKENSGYYU OF TURNAV STANDARS ST					-0		
TPDLTCAVSIHSTVPGLLTSVQTSTG  TPDLTCAVSIHSTVPGLLTSVQTSTG  TPDLTCAVSIHSTVPGLLTSVQTSTG  FRIKSYRKGSYRCIVSEWIAEQGNWQEIQEK AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHSPAPQFPECGFYGLYDKILLFKHDPT SANLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHIGGREAEMRELIQITFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SGGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							
1566	[ ]						
AVEVATVVIQPTVLRAAVPKNVSVAEGKELD LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPTYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFPQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRL VQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRASLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSEIHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	216	1566	A	2477	1	414	
LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHSPAPPECGFYGLYDKILFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS		100,0			_		
RVLARLDRDFLVHSSPHVALSHVDARSYHLL VRDVSKENSGYYY  217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEFASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	<b>!</b>	'					
VRDVSKENSGYYY							
217 1567 A 2480 2 460 CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFPQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	J i						
MQVLVCQHECVRELATRPGRLSPIENFLPLHY DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	217	1567	A	2480	2	460	
DYLQFAYYRVGEYVKALECAKAYLLCHPDD EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAÆGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							
EDVLDNVDYYESLLDDSIDPASIEAREDLTMF VKRHKLESELIKSAAEGLGXSYTEPNYW  218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS					i		DIR ODLIGORIOGRAPILI DOLIVISTI GEDDE
218 1568 A 2483 140 383 AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQRPHALTVHSYRAP 219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS 220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
1568	L	L					VKRHKLESELIKSAAEGLGXSYTEPNYW
SANLLQLVRSSGDIQEGDLVEVVLSASATFED LQIRPHALTVHSYRAP  219 1569 A 2489 3 428 SSRLVLLAGAAALASGSQGDREPVYRDCVLQ CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
LQIRPHALTVHSYRAP							SANLLQLVRSSGDIQEGDLVEVVLSASATFED
CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							
CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	219	1569	A	2489	3	428	
DCKYECMWVTVGLYLQEGHKVPQFHGKWP FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
FSRFLFFQEPASAVASFLNGLASLVMLCRYRT FVPASSPMYHTCVAFAWVS  220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	•		ŀ		İ		
220 1570 A 2498 1 1297 MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	(	}	l		ľ	ľ	
APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS							
APKIPRLVQATPAFMAVTLVFSLVTLFVVDH HHFGREAEMRELIQTFKGHMENSSAWVVEIQ MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
MLKCRVDNVNSQLQVLGDHLGNTNADIQMV KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS		ì	l	ì	1		
KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS			1			·	HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
KEDLEKADALTFQTLNFLKSSLENTSIELHVL SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS	] ]		1	1			MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
SRGLENANSEIQMLNASLETANTQAQLANSS LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS		Ì	j			l	
LKNANAEIYVLRGHLDSVNDLRTQNQVLRNS			İ				
		ļ	J	}		j	
LEGANAEIOGLKENLONTNALNSOTOAFIKSS		1					
	L				l		LEGANAEIQGLKENLQNTNALNSQTQAFIKSS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence	l	09/496 914	ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uence			914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	- Language	/=possible nucleotide deletion, \=possible
		j		sequence		nucleotide insertion
						FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
						VTAQTQKANGRLDQTDTQIQVFKSEMENVN
					ļ	TLNAQIQVLNGHMKNASREIQTLKQGMKNA
			<b>i</b> .			SALTSQTQMLDSNLQKASAEIQRLRGDLENT
<u></u>				L		KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV
						TDEDTRHLSRKFKDWAYGPVYSSLYDLSSLD
			i i			TCGEEASVLEILVYNSKIENRHEMLAVEPINE
						LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
		Į.				AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT GVLFFFTN
222	1572	A	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
	. 13/2	<b>^</b>	2300	3	333	ARHVRDKEEEVDLVMQKVESLRQELRRTER
			]			AKKELEVHTEALAAEASKDRKLREOSEHYSK
			[	ľ		QLENELEGLKQKQISYSPGVCSIEHQQEITKL
						KTDLEKKS
223	1573	A	2544	2	412	NDPAIISNESAAVVHTIVNETLESMTSLEVTK
						MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
			ŀ	0		EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
		İ				KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
						LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
1						CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
						QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
		{				TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
225	1575	Ā	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
	1373	<b>^</b>	2505	124	1	GLGEEEKEAGKKKKKQEEKEKGAVYSR
			i i			VARICKNDMGGSORVLEKHWTSFLKARLNC
1						SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
1						QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP
	'					DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
ł i	,	•				TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
						KTRVRYRLTAISVDHSAGPYH
226	1576	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
						TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
						AGAACDRGMSLEACEAVTRKANRRTYTMG
						VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
	15.1	4.	25,5	-		EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
						ELWMKAMLDAALVQTEPVKRVDKITSENAP
						TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
.			}			KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
						PSEYESGSACPAQTVHYRPINLSSSENKIVNVS
						LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
					1	LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
						HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
						STHDKTLGPGAEEKRRSMRDDTMWQLYEW
1					ŀ	QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
						MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
228	1578		2583	3	330	QRGDVTIDRRHRAHHPKVK  LPFLGLGSVLPQGMVMASPEMNPTICSVFEA
220	17/0	Λ	2,703	ا ا	٠ ،	HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
					.	KIHAKVRGTWPFISPEVRKEGGLPOTGRELLD
[ [			[			PTMGIKPHLWWVAA
229	1579	A	2589	1	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
				!		ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
l				l	-	GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

		tted end   Amino acid sequence (A=Alanine C=Cysteine,
NO: of NO: of hod ID NO:	beginning nucle	
nucl- peptide in	nucleotide location	
eotide seq- USSN		ponding I=Isoleucine, K=Lysine, L=Leucine,
seq- uence 09/496		amino M=Methionine, N=Asparagine, P=Proline,
uence 914		esidue Q=Glutamine, R=Arginine, S=Serine,
	amino acid of per	
	residue of seque	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1     1	sequence	nucleotide insertion
	sequence	GGTCLAGNLCTCPYGFVGPRCETMVCNRHC
		ENGGQCLTPDICQCKPGWYGPTCSTA
230 1580 A 2593	2 138	AVTFSVVFAYVADITQEHERSMAYGLVCMFI
250 1500 11 2555	2   150	LYLLYLLRNAFFLR
231 1581 A 2595	185 2	SGPYTDFTPWPTEEOKLLEOALKTYPVNPPER
		WEKIAEAVPGRTKKACIKRYKVADLRISK
232 1582 A 2596	1 391	STVTGQPRRLLDTAGHQQPFLELKIRANEPGA
		GRARRTPTCEPATPLCCRRDHYVNFOELGW
		RDWILLPEGYQLNYCSGQCPTHLAGSPGIAAS
		FHSAVFSLLKANNPWPGRTSWCVPTARRPLS
		LLYL
233 1583 A 2601	184 403	LLFSDEIIMAAPLRIADVTSGLIGGEDGRVYV
		YNGKETTLGDMTGKCKSWITPCPEEKVNVLQ
		NSIPYWERIT
234 1584 A 2614	178   335	PLTLCLPENNKPPQADAVPDKELTLPVDSTTL
		DGSKSSDDQKIISYLWEKTQ
235 1585 A 2616	2 896	DVLEVYGTGVASTRHEMGTLDKHKELEDLV
		AKFLNVEAAMVFGMGFATNSMNIPALVGKG
1 1 1 1		CLILRDEVNHTSLVLGARLLGATIGIFKHNYA
		QSLEKLLRDAVIYGQPRTRRAWKKILILVEGV
		YSMEGSIVHLPQIIALKKKYKAYLYIDEAHSI
		GAVGPTGRGVTEFFGLDPHEVDVLMGTFTKS
1		FGASGGYIAGRKARILSPPACLVPNTGSHSLH
		RLTRDLQMNEAMVALVTDRLQGWNSGEGN WDRADKFGDLVDYLRVHSHSAVYASSMSPPI
		AEQUIRSLKLIMGLDGTTQ
236 1586 A 2621	1 392	NTSSFPAQPSSPARPSLPHLSQHPSNPLLPLAS
	.   3,2	ADHPQCGRFLPLHEPEPLCPSPSLSYPTLVSS
		WSSPFSSHHGCPPGLYPFPTSPKTIQPPGLAQL
	i	KMLCIPPGRQQLRGAQSMPGHGALSPLLLPP
		A
237 1587 A 2628	398 1	DLVCKISGFGRGPRDRSEAVYTTMSGRSPAL
1 1 1 1		WAAPETLQFGHFSSASDVWSFGIIMWEVMAF
		GERPYWDMSGQDVIKAVEDGFRLPPPRNCPN
1		LMHRLMLDCWQKDPGERPRFSQIHSILSKMV
		QDPEPPNV
238 1588 A 2631	1 1104	WSPCSLTCGVGLQTRDVFCSHLLSREMNETV
		ILADELCRQPKPSTVQACNRFNCPPAWYPAQ
	1	WQPCSRTCGGGVQKREVLCKQRMADGSFLE
		LPETFCSASKPACQQACKKDDCPSEWLLSDW
		TECSTSCGEGTQTRSAICRKMLKTGLSTVVNS
	j .	TLCPPLPFSSSIRPCMLATCARPGRPSTKHSPHI
		AAARKVYIQTRRQRKLHFVGGGFAYLLPKTA
;		VVLRCPARRVRKPLITWEKDGQHLISSTHVT
	1	VAPFGYLKIHRLKPSDAGVYTCSAGPAREHF
		VIKLIGGNRKLVARPLSPRSEEEVLAGRKGGP
	1	KEALQTHKHQNGIFSNGSKAEKRGLAANPGS   RYDDLVSRLLEQGAPCSSSKKKN
239 1589 A 2636	678	MKPDNILLDEHGHVHITDFNIAAMLPRETQIT
2030 A 2030	'   "	TMAGTKPYMAPEMFSSRKGAGYSFAVDWW
	1	SLGVTAYELLRGRRPYHIRSSTSSKEIVHTFET
		TVVTYPSAWSQEMVSLLKKLLEPNPDQRFSQ
	1	
		LSDVONFPYMNDINWDAVFORRI IPGFIPNIK
		LSDVQNFPYMNDINWDAVFQKRLIPGFIPNK GRLNCDPTFELEEMILESKPLHKKKKRLAKK
		GRLNCDPTFELEEMILESKPLHKKKKRLAKK
240 1590 A 2639	389 3	GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIIINRE KVNRDCI ELLDPTTPMRTKCIELLYAALTSSSTDQPKAD
240 1590 A 2639	389 3	GRLNCDPTFELEEMILESKPLHKKKKRLAKK EKDMRKCDSSQTCLLQEHLDSVQKEFIINRE KVNRDCI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1591	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD TCDLCGYNQKLYPCWETQVGQEMYKLMIFD FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	A	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT STTAMVMPIVEAVLQELVSAEDEQLVAGNSN TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL VSFINFFTSVLATLVVFAVLGFKANVINEKCIT QNSETV
244	1594	A	2650	1		MTTTLIGLLKTARLLRLVRVARKLDRYSEYG AAVLMLLMCIFALIAHWLACIWYAIGNVERP YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA WTYTNGIDMNMVTNGTCSSCTSDDGHFILVS NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS TSDSNLNKYSTINKIPQLTLNPSEVKTEKKNSS PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE VVSDPASV
245	1595	A	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP AKPRLQRGGSSASLHNSLMRNSIFQNKIHTLD PHV
246	1596	A	2660	200	506	VLVLQMNYYQMLIIYYVLFFKVNEFLAFEGPI LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS NYFLS
247	1597	A	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF VSLSYLEIIFDPAQLCDSSEHIIS
248	1598	A	2687	1	404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD ENDAVAEESRQVLTICAQFLKWKLPREVYSK DPWHIKPTEAGTICRFFEKKCKGKINILEQTL MYSKNPKL
249	1599	Ā	2692	1	440	FRRRRRRERDCAAQGARRHCRHLAECKLV SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK AIDLSRNQFQDFPEQLTALPALETINLEENEIV DVPVEKLAAMPALRSINL
250	1600	A	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	A	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	į.	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l	Ì	ł	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
<b></b>		<del> </del>	<b>├</b>	sequence		nucleotide insertion
i				1		QKRKKKAPDHSSGRKEELVTTHTVDKLETKK PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
l	}	1	]	ŀ		RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
1	<b>\</b>	l		Ì		GAGDCL
252	1602	A	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG
			}		-	TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
1		1				VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
	İ	ļ				MNVTGPQIRDGGVYRCTARNLVGSAEYOARI
<u> </u>		Ì				NVRGPPSIRAMRNIT
253	1603	Α	2698	65	401	ACCQWRRTLIPAKSTTVSCTISTPHHPFRGSYS
ł		l				FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL
Ì						PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
						AAFTNNIASSTIIL
254	1604	A	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
200	1605	<u> </u>				RVRGPWEAGPGVGY
255	1605	A	2700	1	842	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
						AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
1					ļ	DKETLENMMQRHEEEAHEKGKILSEQKAMIN
		1	ĺ			AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
						NMKAQEEMISELRQQKFYLETQAGKLEAQN
						RKLEEQLEKISHQDHSDKNRLLELETRLREVS
						LEHEEQKLELKRQLTELQLSLQERESQLTALQ AARAALESQLRQAKTELEETTAEAEEEIQALT
						VGLGSNIFRLLKASARMSVELALSILAHP
256	1606	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
				_		LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL
]						YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA
1						FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
						PYEAARMFFEGLR
257	1607	Α	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
						FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
ľ ,						NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
						DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI
1250	1600					GGNDL
258	1608	A	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
	.			1		MTRKELLTVYSSEDGSEEFETIVLKALVKACG
						SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
						QWRSFHLEASLMDALLNDRPEFVRLLISHGLS LGHFLTPMRLAQLYSAAPSNSLIRNLLDOASH
{					*	SAGTKAPALKGGAAELRPPDVGHVLRMLLG
						KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
			1		·	KATSPLSLDAGLGQAPWSDLLLWALLLNRA
]			}			QMAMYFWEMGSNAVSSALGACLLLRVMAR
		l	.	İ		LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
	1					SSEVRAARLLLRRCPLWGDATCLQLAMQAD
						ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	Α	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
	-	l		.		GGFSQREMVTGERSPSPEEEEEEEEGFGERA
1 1		i				SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
1 1	1	1	1			PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
L	1616		2525		485	GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
			İ			LEYLQRYHIIHRDIKPDNILLDEHGHVHITDFN
	l	- 1	1		. 1	IATVVKGAERASSMAGTKPYMAPEVFQVYM
1 1	İ	}			į	DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
	1		1	J		HSVTPIDEILNMFKVERVHYSSTWCKGMVAL LRK
261	1611	Ā	2730	3	547	
		^	2130	- I	371 ·	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH
						PYTPGOCKYT FASIBLIADSPLEWA I JETVIAKIH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1.		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	Soquanoc	/=possible nucleotide deletion, \=possible
		1		sequence	[	nucleotide insertion
<b></b>			<b>-</b>	Soquette	<del></del>	RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
1					ļ	RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
		!			3	DIFVDRRVSALAVVNECGTHPQDERLGLGW
i I		i			Ī	CI CEDCCEERI EDA AFTOR
262	1612	A	2733	3	431	GLGEPGSEERLFPAAITSR
202	1012	^	2/33	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
1		1				GRLVKLSLANNNLVGVHEDAFETLESLQVLE
		ł				LNDNNLRSLSVAALAALPALRSLRLDGNPWL
]						CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
362	1613		2007		0.40	ESRRISLRACRRPASRV ·
263	1613	Α	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF
1	ļ	1				LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
1		ŀ				ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
						RLISPLVNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV
}						DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
						DTVDVLDPPEDSGKQVDL
265	1615	A	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
}						LAAVETTVLVLIFAVSLLGNVCALVLVARRR
1 1						RRGATACLVLNLFCADLLFISAIPLVLAVRWT
						EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
						SLER
266	1616	A	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
1 1						LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
1 1					;	V
267	1617	A	2760	434	714	ASRLEKONSTPESDYDNTPNDMEPDGMGYM
1			2.00			HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
						LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	A	2762	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
			1	-		AVLLLLLLSLALGLVLAALGLFVHHRDSPL
						VQASGGPLACFGLVCLGLVCLSVLLFPGOPSP
						ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
					-	LPLSWAE
269	1619	A	2772	3	243	
120	1017	^	21/2	,	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
1 1		j	}			LSKNLSFSEFCFDVSY
270	1620	A	2789	1	486	
2,0	1020	n	2107	*	400	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
. [		(		i		VEQIAKAEETHSSLSQELQARLQTVTREKEEL
			. [		ļ	LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
	1	- 1	1		1	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
		İ	ł			KAYDELRLQSEAFKKHSLDLLSKERELNGKL
271	1601		2505			RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
	}		ĺ		ļ	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
		1	[	.		RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
	ľ	- 1	1		ľ	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
		l		1		VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
		1				SQRDDGKGALNIVEFLRKREELHQILSTVKQP
272	1622	A	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
		1	1			RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
ĺ	ĺ	ĺ		Í	Į.	GYRKVVSNNCTDGVREQYTAKPOKCPGKAP
	ŀ				ļ	RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
		-	1	.4.		VORTLIOVDFGDGIAVSYVNLSSMEDGIXHV
		İ				YONXGIXRXTVQVDNSLGS
273	1623	A	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
*						DEEGONOLGIRIHODIPLPPRRRELPALRTING
	- 1			}		KADSLNVSRNSVMQELSELEKQIQVIRQELQL
	- 1	- 1	ĺ	1	Ì	AVSRKTELEEYH
274	1624	A	2805	168	320	ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE
	.024	-	_555	• •	.	
	l					IFIARNGVVGETLTHCKRV

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghtamic Acid,
nucl-	peptide	~~~	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
acate.		l	] 7.7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ł		peptide	Sequence	/=possible nucleotide deletion, \=possible
i I						nucleotide insertion
275	1625	<del>  </del>	2812	sequence	321	
2/3	1023	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
276	1626	<del></del>	2012	4.		MGKJIFQ
276	1020	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
		}				KYQQQTVVAIDLAGDETIPGSSLLPGHVQAY
J						QVGPVRRNGEAGPG
277	1627	Α	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
1.						LFISYLHTPKHKQHEVLQAMGSILGITGEEME
l			1		1	PLFQEEHGTATRWMTGWLEGGSKSVPKTPL
						GLNQQPALNGSFSELFVKFLKTESLSSTLPTX
			L			LPPHNSPGKIK
278	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
						VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
					Ì	PVDQNPRLV
279	1629	Α	2822	342	1	PLIPANLPAHSNPLQPLPSLPHPFLPATHKFPT
						TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
						CRLPSRPSVSRELPPQSGPASSVPLAPTPLPDS
						VPSORHPTXPPPAS
280	1630	A	2825	307	77	PSMVWSYHWGVKOKRLALCVFSFEEGGRRK
1		1.	1025	307	''	CGOYWPLEKDSRIRFGFLTVTNLTGAVGEPG
						VAFQCDGQRRREPTC
281	1631	A	2827	81	381	KMGTAVWVPKEKEKRDKASOEGGDVLGAR
201	1031	^	2021	01	301	
1						QDCTPSLKSLVATGNLLDLEETAKAPLSTVSA NTTNMDEVPRPQALSGSSVVWVSGCVASRS
				,		
282	1632		2020	471	160	VILSLTSG
282	1032	A	2830	471	160	KLPXDKYELEPSPLTQYILERKSPHTCWQVFV
						TSSGKYNELGYPFGYLKASTTLTCVNLFVMP
						YNYPVLLPLLDDLFKVHKLKPNLKWRQAFDS
<u></u>						YLKTLPPYYL
283	1633	A	2835	462	148	VSPALSLTPTIFSYSPSPGLSPFTSSSCFSFNPEE
1 1						MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
						PLQCQMHPEESTQFSIKLQPPPVGRKNRERVE
		_				SSEESAP
284	1634	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
						DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
[ ]	ļ					KSLAETVLNFPLDKSLLLRCSNWDAETLTED
} }						QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
L						KR
285	1635	A	2843	20	271	PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL
						VCDRVSEDGINRQQAQEWCIKHGFELVELSP
]						EELPEEDGKCLCVRRKYGTYI
286	1636	Ā	2845	197	278	TAEDVLTVAYEHGVNLFDTAEVYAAGK
287	1637	Â	2851	2	427	FVAEVRREWAKYMEVHEKASFTNSELHRAM
<del> '</del>				-	· <del>-</del> ·	NLHVGNLRLLSGPLDQVRAALPTPALSPKDK
j j				ļ	1	AVLQNLKRILAKVQEMRDQRVSLEQQLRELI
	[	ĺ	ĺ	1		QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL
			1	l		
200	1620	<del></del> -	2050	<del></del> _	460	KVYLEQNLAAQDRVLCALT
288	1638	A	2859	2	469	FVNLGILTCIECSGIHREMGAHISRIQSLELDK
	` <u> </u>			ļ	l	LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP
				I		TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA
l l	ľ	- 1	1	l	i	KLNELLEAIKSRDLLALIQVYAEGVELMEPLL
						EPGQELAETALHLAVRTADQTSLHLVE
289	1639	A	2861	2	454	FVASGGPATARMSDSQFFCVAEERSGHCAVV
	. )	ļ	j		j	DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI
·	1	l	ł			DSGLWRMHLMEGELPASMSGSCGACINGKL
	ł	ł	}		ļ	YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK
	!			ļ		ITDFEGQPPTPRDKLSCWVYKDRLIYFG
290	1640	A	2868	1 -	378	FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI
				-	'	SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF
	1	l	j	i		PDCASCLQAQDPLCGWCVLQGRCTRKGQCG
						- DO NOT TO THE POST OF THE PO

NO. of continue   No. of con	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decide   Seq			1				
Decided   Deci			1				
United   914							
maino seid residue of peptide sequence   Filtmonine, V-Valine, W-Tryptophan, Y-Sup codou, peptide sequence   Sequence		uance	1				
Personage   Pers	uence		l	914			
peptide   peptide   peptide   peptide   peptide   peptide   pert			l				
	1		l			sequence	
RAGOLNOWLWSYEEDSHCLHIQSLLPGHHPR   QE	ŀ		1	1	peptide		/=possible nucleotide deletion, \=possible
RAGOLNOWLWSYEEDSHCLHIQSLLPGHHPR   QE			1	ĺ	sequence	1	nucleotide insertion
1641				<del>]                                    </del>	<del></del>		RAGOLNOWLWSYEEDSHCLHIOSLLPGHHPR
1641							OE.
PFTPKSIRSHPÖHNYTVÄLVÄNNÄCTENÄCYSÄ   GVSRSKDVPPEGPPTRGVTTPKSÄ VPROPELL AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKVINAENAAHKSEKFRAMATRTRQEYLKD     AKRIEGE VEEEDEN KVEETEKEREEDIST     AKRIEGE VEEEDEN KVEETEKEREEDIST     AKRIEGE VEEEDEN KVEETEKEREEDIST     AKRIEGE VEEEDEN KVEETEKEREEDIST     AGATHPDAAPSGERTCGSEGSRSVLDLVYSLSG     GRCYLLITLIARSFDLRTMRRKKLDDVSIPLL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILPLAGGRQAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDL     LILLAGGRGAYDH     LILLAGGRG	291	1641	Δ	2870	1	395	
	291	1041	\ ^	2870	1	202	
AKVINAENA-AHKSEKFRAMATRTRQEYLKD   LA				1		1	
LA	1				İ		
1642			1			1	1
PPPPA MYPYSPRYVAVHCHGML VSCWCHI	<u></u>						
1643	292	1642	Α	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
1643							
CAGTHPDAAIPSGERTCGSEGSRSVLDLYNYF	293	1643	A	2878	7	427	
SEPEKLTAENRYYCESCASLQDAEKVVELSQ   GPCYLLTLLERSPOLRTMRRRKILDVSIPLL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGQAYDL   RLPLAGGRGGAYDL   RLPL	-50		1	20,0	•	T/	
GPCYLLITLLRFSSPLATMRRRKILDDVSIPIL   LRLPLAGGROGAYDL	l. I		l				
	1		1				
1644	i		i				
1645						·	
295	294	1644	Α	2879	109	245	
NNCVGEQNIRFFCALHCKSKHFCIEFTLNTNF   FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS   LSESISQ							
NNCVGEQNIRFECALHCKSKHFCIEFTLNTINF	295	1645	Α	2880	3	320	LASSOHGILNNLSLLFSICKTCIRTMDHHCPRA
PRICELPGÄEKSTIDAPPSLQPFLQDSKYNTALS	1 1		i	1			
LSESISQ							
296				}			
RLQEFSQKMDQVRGHWPVST	206	1646		2002	200	262	
1647	250	1040	^	2092	209	303	
KLYSTMGRPLRDRKNPACREMAVVLLANLA   QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT   QIQQSQASLLHMINPPFEPTISVDMMRRACRA   LLALAKVDDNHSEF							
QGDSLAARALAVQKGSIGHLLGFLEDSLAAT   QIQQSQASLLHMHNPPFEPTSVDMMRRACRA   LLALAKVDDMHSEF	297	1647	Α	2893	8	424	
1648							KLYSTMGRFLRDRKNPACREMAVVLLANLA
LLALAKVDDNISEF	1						QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT
LLALAKVDDNISEF	1				İ		
1648	i i						
SGLLNASAQVNL	298	1648	Ā	2894	310	445	
1649		10.0		20,1	310	113	
GYFQAYNVLILTMQASLPKVLRFCACAGMIY   LGYTFCGWIVLGPYHDKFENLNTVAECLFSL   VNGDDMFATFAQIQQKSILVWLFSRLY1LYSFI   SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF	200	1640		2000		400	
LGYTFCGWIVLGPYHDKFENLNTVAECLFSL	299	1049	A	2090	1	492	
VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI   SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF							
SLFTYMILSLFIALITDSYDTIKKFQQNGFPETD   LQEF							
LQEF	1						
1650	i i			]			SLFTYMILSLFIALITDSYDTIKKFQQNGFPETD
1650							
TVTVRFVNKADFPKVRAKEQTFMFPENQPVS   SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ   LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP   FSSYEKLDITVLDVNDNAPIF	300	1650	A	2901	1	445	
SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ						·-	
LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP   FSSYEKLDITVLDVNDNAPIF							
FSSYEKLDITVLDVNDNAPIF		. 1			l	İ	
1651							
EPCGWMYDHAKWLRTTWASSSSPNDRTFPG KPAVSEDMKELRPACSTYFNPRFPYKL  302 1652 A 2909 2 412 GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	201	166:		2022			
SPAVSEDMKELRPACSTYFNPRFPYKL	301	1001	Α	2902	162	433	
302   1652   A   2909   2   412   GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV   EENDHQDLVSVHIYVTQLAEKFDLRITMLYI   CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN   SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ   LVNRQDRAHFM     303   1653   A   2914   291   453   KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE   VPPTSILEHLQRRKIMKRPSSCS   VPPTSILEHLQRRKIMKRPSSCS   VPPTSILEHLQRRKIMKRPSSCS   304   1654   A   2926   179   354   PGVPSQALRKAESLKKCLSVMEAKVKAQTAP   NKDVQREIADLGEVGAASLPPSSGPGA   305   1655   A   2938   135   438   GMGYLHAKGILHKDLKSKNVFYDNGKVVIT   DFGLFSISGVLQAGRREDKLRIQNGWLCHLA   PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE   LHAREWP   VRWNSCVNCSCAFGNGASLSTSLGESSGCLW   EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		Ì			ŀ		EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
302   1652   A   2909   2   412   GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV   EENDHQDLVSVHIYVTQLAEKFDLRITMLYI   CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN   SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ   LVNRQDRAHFM     303   1653   A   2914   291   453   KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE   VPPTSILEHLQRRKIMKRPSSCS   VPPTSILEHLQRRKIMKRPSSCS   VPPTSILEHLQRRKIMKRPSSCS   304   1654   A   2926   179   354   PGVPSQALRKAESLKKCLSVMEAKVKAQTAP   NKDVQREIADLGEVGAASLPPSSGPGA   305   1655   A   2938   135   438   GMGYLHAKGILHKDLKSKNVFYDNGKVVIT   DFGLFSISGVLQAGRREDKLRIQNGWLCHLA   PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE   LHAREWP   VRWNSCVNCSCAFGNGASLSTSLGESSGCLW   EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA						j	KPAVSEDMKELRPACSTYFNPRFPYKL
BENDHQDLVSVHIYVTQLAEKFDLRTTMLYI CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS  304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	302	1652	Α	2909	2	412	
CERHFQKVLNRSLFTGLRSITHFGRPPFEPFFN SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS 304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	ŀ		}				
SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ LVNRQDRAHFM  303 1653 A 2914 291 453 KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS  304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA  305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		-	1	l	ļ	l	
LVNRQDRAHFM   303   1653   A   2914   291   453   KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE   VPPTSILEHLQRRKIMKRPSSCS   304   1654   A   2926   179   354   PGVPSQALRKAESLKKCLSVMEAKVKAQTAP   NKDVQREIADLGEVGAASLPPSSGPGA   305   1655   A   2938   135   438   GMGYLHAKGILHKDLKSKNVFYDNGKVVIT   DFGLFSISGVLQAGRREDKLRIQNGWLCHLA   PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE   LHAREWP   2   329   VRWNSCVNCSCAFGNGASLSTSLGESSGCLW   EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	ľ	ì	ľ		İ	ľ	
303   1653   A   2914   291   453   KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE VPPTSILEHLQRRKIMKRPSSCS   304   1654   A   2926   179   354   PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA   305   1655   A   2938   135   438   GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP   306   1656   A   2944   2   329   VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA			l				
VPPTSILEHLQRRKIMKRPSSCS  304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA  305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	200	1655					
304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	303	1653	A	2914	291	453	
304 1654 A 2926 179 354 PGVPSQALRKAESLKKCLSVMEAKVKAQTAP NKDVQREIADLGEVGAASLPPSSGPGA 305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP 306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		i					VPPTSILEHLQRRKIMKRPSSCS
NKDVQREIADLGEVGAASLPPSSGPGA  NKDVQREIADLGEVGAASLPPSSGPGA  SOS 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	304	1654	A	2926	179	354	
305 1655 A 2938 135 438 GMGYLHAKGILHKDLKSKNVFYDNGKVVIT DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA			ł				
DFGLFSISGVLQAGRREDKLRIQNGWLCHLA PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	305	1655	A	2938	135	438	
PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE LHAREWP  306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA						.50	
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA		1		1	1	1	
306 1656 A 2944 2 329 VRWNSCVNCSCAFGNGASLSTSLGESSGCLW EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	1	ł	i	ľ	1		
EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA	200	1.7.5.2					
	305	1656	A	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
	. ]	1	]	J		3	
I I I I I I I I I I I I I I I I I I I			- 1				LTKGALWAVFLLAGSALLCAEVTGVIWROPE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	I	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
<u></u>			<u> </u>	sequence	<u> </u>	nucleotide insertion
	172-	<b> </b>	1	L		SKTKLSFKVSSSA
307	1657	A	2950	2	411	NYLCIAKNSAGSAMGKTRLVVQVPPVIENGL
1		l		ŀ	1	PDLSTTEGSHAFLPCKARGSPEPNITWDKDGQ
1		l	[ :	[	[	PVSGAEGKFTIQPSGELLVKNLEGQDAGTYT
		۱.	[			CTAENAVGRARRRVHLTILVLPVFTTLPGDRS
308	1658	A	2951	1	407	LRLGDRLWLR  DTBDDDVDEDNEEDAESODVBTTSVSVAEDA
300	1030	^	2771		70/	PTRPPRVRFDNEFDAESQRKRTTSVSKMERM DSSLPEEEEDEDKEAINGSGNAENRERHSESS
1				!	1	DWMKTVPSYNQTNSSMDFRNYMMRDETLEP
			]		[	LPKNWEMAYTOTGMIYFIDHNTKTTTWLDP
1 !						RLCKKAKAPEDC
309	1659	A	2954	2	179	QDFLTLTLTEPTGLLYVGAREALFAFSMEALE
"		· ·	2,54	ļ <b>-</b>	```	LQGAVRGGAVGGSRACQRARPRGAVLG
310	1660	A	2959	1	419	QDMMERAIIDTFVGHDVVEPGSYVQMFPYPC
				<sup>-</sup>	1	YTRDDFLFVIEHMMPLCMVISWVYSVAMTIO
1						HIVAEKEHRLKEVMKTMGLNNAVHWVAWFI
						TGFVQLSISVTALTAILKYGQVLMHSHVVIIW
			1			LFLAVYAVATIMFCF
311	1661	A	2963	3	465	MKPQMPGLGAPNGYGPGRGRAGVPGGPERR
						PWVPHLLPFSSPGYLGVMKAQKPGAGEGMK
1			[	1	·	PQKPGLRGTLKPQKSGHGHENGPWPGPCNA
				1		RVAPMLLPRLPTPGVPSDKEGGWGLKSQPPS
1						AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ
Li	l	· .	<u> </u>			кі .
312	1662	A	2967	3	405	WLAQEWSPCTVTCGQGLRYRVVLCIDHRGM
1					,	HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV
						EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS
					!	ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI
	1660		2022		426	DECEGPKPA
313	1663	A	2969	2	430	VVADNCRQGYLDALRFLERRGLTKEPVLWT
						LVSKEPPAPADGNWDAGCDQRRKGGLSLNW
				:	1	KVPHVQVKDVPNFEQLSPELEAALKKACTRD
					ļ	PSRWARFWHSGPGQVLTYLLLPCTLPFEYTYF RSRRLVVWLPDVPADLWWMO
314	1664	Ā	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
ا ``` ا			/1			LDALGRGVFVNASGLRLLDLSSNTLRALGRH
j .	ļ	<u>.                                    </u>				DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA
, , 1		İ		ĺ	İ	LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
	l			· .	ļ <b>l</b>	SSNRM
315	1665	Ā	2973	1	525	ITVSTHASGSPFGLEPQSGWLWVRAALDREA
[ [	1	.				QELYILKVMAVSGSKAELGQQTGTATVRVSI
						LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV
] ]		1				FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ
	ſ		ſ	ĺ	' <b>1</b>	TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP
<u> </u>						RSTTGTVHVAVLDLNDNT
316	1666	A	2978	2	400	ELVVELVSAGKSGPERNTYEVQVVTGNVPKA
		ł				GTDANVYLTIYGEEYGDTGERPLKKSDKSNK
		Į	1		' I	FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR
, ·	.		1			AGWFLDRIDITDMNNEITYYFPCQRWLAVEE
احيا	[					DDGQLSRE
317	1667	Ā	2981	3	440	VLNCQGRPTRPVRINGDGQEVLYLAESDNVR
ļ ļ	J	ļ	J		J	LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH
	l	ļ	1			HRENVFLSYQDKRINHGSLPHLQHRVRFAAS
	ļ	1	į		, [	DPSQYDASINLMNLQVSDTATYECRVKKTTM
	1662	<del></del>	0000			ATRKVIVTVQARPAVPMCWTEGQ
318	1668	A	2995	119	414	LPEKEFPIIRKSSSLKVTKCLFTEQPKPIILRFA
1			Ì		İ	ENYDARLLRIDIANTLREQVQELFNKTYGKQ
210	1660	<u> </u>	2000		222	RRTPGEGHVAAVDREVAGFPVPAEGISGETIH
319	1669	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan,
1				peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
-		<del> </del>		Sequence		STLALSHSAQVLASASGRSSTTAHCQIRVWD
	1	İ	[		[	VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
						GDHDGRTLALWGTGHL
320	1670	Α	3000	693	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA
		l				PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
		1	[			ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
						TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	Α	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
	•					GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
					]	GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
322	1672	A	3007	192	447	WNQKRYEALGEIITKYVYELLEKDCNSKKVS
J22	1072	^	3007	192	***/	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
						DGAASPRNVGHNIYILAHOLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPIIASVSEH
		**		•		QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND
						ERVFGKRGF
324	1674	A	3020	523	797 .	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
[ [						YLFIFFEMESHSVTHAGVQRHNLNSLOPLPPG
	_					FKRFSCLCFLSSWNYRGAPPGPANF
325	1675	A	3022	2	156	NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
226	1676	<u> </u>	2000		150	GNHKAFKKELRQCRWQVGAL
326	10/0	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
327	1677	A	3027	1	385	FCLNKEALKDEFE LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
""	10,,	Λ.	3021	•	262	GFYGLYDKILLFKHDPTSANLLQLVRSSGDIO
						EGDLVEVVLSASATFEDFOIRPHALTVHSYRA
						PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
						RC
328	1678	A	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
						LTGALNAHNKAAVDWGWQGLIAYGCHSLV
1	1		ĺ	ſ		VVIDSITAQTLQVLEKHKADVVKVKWAREN
.						YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
1		,				HPPNYIVLWNADTGTKLWKKSYADNILSFSF D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
			]	<i></i>	′	GNIEYKKLVNPSQYRFEHLVTQMKWRLQEG
						RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
		-				HRMAEKVGADITVLREREVDYDSDMPRKITE
		ĺ	ĺ			VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
}						LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
						TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
220	1/00	<del></del>	2046			W
330	1680	A	3040	3	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
			]	l		LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE
	ļ	.		[		WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ VVLTMTNMGPVDTATYYCAQFARGARGSN
	i	- 1		l		WFDPWGO
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK
	-	.		i		MPRGIAIDWVAGNVYWTDSGRDVIEVAOMK
		-		l	j	GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
				l		WGNHPKIETAAMDGTLRETLVQDNIQWPTG
	1	ı	- 1	ĺ		LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
		-	l	1		VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
1			- 1	l	l	PKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
*					· 1	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG KPI DNGTCVPVRSPTPPDD APPROTCAL OCEN
	ľ	į	. 1		ļ	KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
						GGSCFLNARRQPKCRCQPRYTGDKCELDQC

Cross	I orom	134	Tera	Dun at a 1	I November 1	
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	поа	in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide		Į.	USSN	location		F=Phenylalanine, G=Glycine, H=Histidine,
seq-	seq- uence	j	09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
nence	иенсе	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
baile		l	/17	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	<b> </b>	ĺ		peptide	Sequence	/=possible nucleotide deletion, \=possible
		Į	ļ	sequence		nucleotide insertion
	<u> </u>			554451.55		WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC
		Ì	į			TQQVCAGYCANNSTCTVNQGNQPQCRCLPG
		l	Ì			FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
				ĺ	[	CRCTAYFEGSRCEVNKCSRCLEGACVVNKOS
	1				ļ	GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT
		1			İ	MNSKMMPECQCPPHMTGPRCEEHVFSQQQP
		1				GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG
						AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT
		1			i .	LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD
]	j	j				MPKSPFKRRRSMNEIKNLQYLPRTSEPREVLF
						EDRTRAHADHVGQGFDWQSTAAVGVLKAV
				l		QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ
		i				LDLHEPPVSQCVQWVDEAKLNQMRREGIRY
		ľ	1		ĺ	ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW
	١.					HIRLKQYHPVVEATQNTESNSNMDCGLTGKR
						ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP
						PASE
333	1683	Α	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL
		1				YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT
[						ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG
	2-227					CSSKTWKVAPFVRAWWRP
334	1684	Α	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD
						EASGANDEIVQLRSEVDHLRREITEREMQLTS
007	4.600	ļ.,				QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR
			]			NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA
		l	·		•	YNDVQYQGHYYEWLPRYNDPAAPCALKCH
		Ì	ŀ			AQGQNLVVELAPKVLDGTRCNTDSLDMCISG
i			l			ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGOSKSHVSPEKREENVIAVPLGSRSVRI
1		-	1			TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV
	-:	ł	i i			ENTIVEFORGSEROTFKIPGPLMADFIFKTRY
		i				TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT
						CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA
				•	•	LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG
						AWDWRLGSPACPHWGLHKLPRLWDPLSLYP
						VLCWGT
337	1687	Α	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE
						QQQRRHPQHLHQQHHGDAAQHTRTWKLQT
				l		DSNSWDEHVFELVLPKACMVGHVDFKFVLN
				.		SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF
						LEDHKEDILCGPVWLASGLDLSGHAGMLTLT
				l		SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI
						CNGGMRPVVRLPSLKHQSNKGYSLASLLAK
						VAAGKEKSSNVKNENTSGTRK
338	1688	Α	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK
						DKVYGVADSCTSLLSGRNRCKLGLLSLHETIL
						SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE
						EELNP
339	1689	Α	3063	236	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV
						PSSVTTMLSWV
340	1690	Α.	3065	3	1249	DLWQFTPLHEAASKNRVEVCSLLLSYGADPT
j			}			LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL
Į					ĺ	QAAREADVTRIKKHLSLEMVNFKHPQTHETA
					l l	LHCAAASPYPKRKQICELLLRKGANINEKTKE
ı	ı		ĺ		1	M. World T. Court   An Maria Country of Coun
						FLTPLHVASEKAHNDVVEVVVKHEAKVNAL DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide		l in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	]	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine.
"""	İ	1	* * *	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	l				soquence	1-1 ylosine, A-Onknown, 1-Stop codon,
ſ	l	ĺ	ĺ	peptide		/=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion
l .		j				IISLQGFTALQMGNENVQQLLQEGISLGNSEA
	l					DRQLLEAAKAGDVETVKKLCTVQSVNCRDIE
1			[			GRQSTPLHFAAGYNRVSVVEYLLQHGADVH
	ł		i I			AKDKGGLVPLHNACSYGHYEVAELLVKHGA
						VVNVADLWKFTPLHEAAAKGKYEICKLLLO
1	l		J			HGADPTKKNRDGNTPLDLVKDGDTDIQDLLR
		ł				GDAALLDAAKKGCLARVKKLSSPDNVNCRD
						TOGRHSTPLHLAGK
341	1691	A	3070	1	547	GVLIPSFQNQLFADILAGIESVTSEHNYQTLIA
1 3 7 7	1071	<b>'</b>	3070	•	J 341	NYNYDRDSEEESVINLLSYNIDGIILSEKYHTI
						RTVKFLRSATIPVVELMDVQGERLDMEVGFD
			i			NRQAAFDMVCTMLEKRVRHKILYLGSKDDT
	ł	}	1			RDEQRYQGYCDAMMLHNLSPLRMNPRAISSI
						HLRMQLMRDALSANPDLDGVFCTN
342	1692	A	3073	463	3	RINRCRKPSDADILVPGDTISLIGTTSLRIDYNE
						IDDNRVTAEEVDILLREGEKLAPVMAKTRILR
1						AYSGVRPLVASDDDPSGRNVSRGIVLLDHAE
1			,			RDGLDGFITITGGKLMTYRLMAEWATDAVC
						RKLGNTRPCTTADLALPGSQEPAKVP
343	1693	Α	3075	250	1	LLIYLAIFAPVAMSALAGVKSVQQVRIRAAQS
		1			-	LGASRAQVLWFVILPGALPEILTGLRIGLGVG
1	ļ.	l	i			WSTLVAAELIAATRGLGFM
344	1694	A	3076	2	138	LYFDAYLQSLQVAAISTFCCLLIGYPLAWAV
744	1094	^	3070	2	170	AHSKPSTRNILLLL
345	1695		2070	469		
343	1095	Α	3078	469	3	LKIRGQRIELGEIDRVMQALPDVEQAVTHAC
						VINQAAATGGDARQLVGYLVSQSGLPLDTSA
						LQAQLRETLPPHMVPVVLLQLPQLPLIANGKL
						DRKALPLPELKAQAPGRAPKAGSETIIAAAFS
						SLLGCDVQDADADFFALGGHSLLAMKLAT
346	1696	Α	3082	404	2	QNITSKDLDVRLDPQTVPIELEQLVLSFNHMI
						ERIEDVFTRQSNFSADIAHEIRTPITNLITQTEI
						ALSOSRSOKELEDVLYSNLEELTRMAKMVSD
[ '			1			MLFLAQADNNQLIPEKKMLNLAHEVGKVFD
						OFEALPE
347	1697	A	3084	3	340	NELTFKEAEISKLYTKVHPAYRTLLEKROALE
"	1057	**	500.	-	370	DEKAKLNGRVTAMPKTQQEIVRLTRDVESGQ
						QVYMQLLNKEQELKITEASTVGDVRIVDPAIT
240	1.000		2006	702	10	QPGVLKPKKGLIILGAI
348	1698	A	3086	723	10	TQAMVWQQKACAEDDPQLSGRHWLHAATL
						YNIAAYPHLKGDDLAEQAQALSNRAYEEAA
]				l		QRLPGTMRQMEFTVPGGAPITGFLHMPKGDG
						PFPTVLMCGGLDAMQTDYYSLYERYFAPRGI
						AMLTIDMPSVGFSSKWKLTQDSSLLHQHVLK
]						ALPNVPWVDHTRVAAFGFRFGANVAVRLAY
					Į	LESPRLKAVACLGPVVHTLLSGLKCQQQVPE
1 '						MYLDVLASRLGMHDASTKSSTRENH
349	1699	Α	3087	2	249	RIRSSDPEITLAGTPLHAAYLIGMTLICAGFSV
						GFGVAMSOALGPFSLRAGVASSTLGIAOVCG
				j	l	SSLWIWLAAVVGIGAWNM
350	1700	A	3099	3	424	EAPEATPOPSOPGPSSPISLSAEEENAEGEVSR
	1700	Λ.	3077	·	744	
1			i	ļ		ANTPDSDITEKTEDSSVPETPDNERKASISYFK
} I						NQRGIQYIDLSSDSEDVVSPNCSNTVQEKTFN
			1			KDTVIIVSEPSEDEESQGLPTMARRNDDISELE
						DLSGMEDLK
351	1701	A	3108	2	404	IKKNHIIGYQLLHRRALFEKRTRLSDYALIFG
		ļ	í			MFGIVVMVIETELSWGAYYKAPLYSLALKCL
<u> </u>	Ì	İ	İ			ISLFTIILLGLTIVYHAREIQLFMANYGADDWR
] [		ŀ		-		SALTYEPIFLILLEALRGVIHATPCRVSLSLWD
	ł	1	ł	ì		GLDLP

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methiomine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
352	1702	A	3110	341	2	AQLAEVCPPQTLLTTNTSSISITAIAAEIKNPER VAGLHFFNPAPVMKLVEVVSGLATAAEVVE QLCELTLSWGKQPVRCHSTPGFIVNRVARPY YSEAWRALEEQVAAPEVI
353	1703	A	3111	3	188	HFSLFRIAFAVFLTYMTVGLPLPVIPLFVHHEL GYGNTMVGIAVGIQFLATVLTRGYAGRLA
354	1704	A	3116	367	225	WQLFHLNGTFLNIGETDTESCVNGWVYDRSS FPFSNMTEVRGLVFLS
355	1705	A	3117	101	53	VINLVYLISSPRPELKPVDKESEVVMKFPDGF EKFSPPILQLDEVDFYYDPKHVIFSRLSVSADL ESRICVVGENGAGKSTMLKLLLGDL\APVRGI RHAHRNLKIGYFSQHHVGAAGT*TFSACGNL LGTQVFLGRPEEEYNRHQLGFGMGISGELGHA SSLPACLGGQKEAEVAFCSDGLLPCPNFL\IL\ DEPTNHLGHGRAIEALGPCLQTISGVGVILVS HE*SALSRLVCRELLWVC*GRSTSPF
356	1706	A	3121	137	466	RGGRDWGEHNQRLEEHQARAWQGAMDAG AASREHARWQGTGLAPGTRVAVAPTCVQGL PQERSVCRPFFSSRWREGPVWALGAGAHGKP RWSGGVRCVVRGGRWFTPAPH
357	1707		3124	1249	229	MLEAPGPSDGCELSNPSASRVSCAGQMLEVQ PGLYFGGAAAVAEPDHLREAGITAVLTVDSE EPSFKAGPGVEDLWRLFVPALDKPETDLLSH LDRCVAFIGQARAEGRAVLVHCHAGVSRSV AIITAFLMKTDQLPFEKAYEKLQILKPEAKMN EGFEWQLKLYQAMGYEVDTSSAIYKQYRLQ KVTEKYPELQNLPQELFAVDPTTVSQGLKDE VLYKCRKCRRSLFRSSSILDHREGSGPIAFAH KRMTPSSMLTTGRQAQCTSYPIEPVQWMESA LLGVMDGQLLCPKCSAKLGSFNWYGEQCSC GRWITPAFQIHKNRVDEMKILPVLGSQTGKI
358	1708 .	Α	3127	816	139	EVETLGPRTPGP/EAQSPTPGSCPGWQEPSPGP TPPP*LSGPGPQGAPVLGKLLPDPEETPAGKTP LGKHFWWGLIPVTSANFSPGAAA*FGGALSPP GGDL/GHMLLQGPPSFFRLQQQ*QTPPGSHSP PTANREINPGPAAAADTRSCWGHKRSWRGW RGLAPWRLGFGSPGIP*PAPAGIP/GRPTWEGG KGAGGKPSETLTRSPPVWRGKRGSANGFLSW VQILQ
359	1709	A	3132	3	191	HEHLLLLLCVFLVKSQGVNDNEEGFFSARG HRPLDKKREDAPNLRPALADUTVCDYRAQIA *AASTPKRAASIAHNAVSCR*AQIA
360	1710	A	3134	1	286	REPPRPALLFF*DRVSLCCPGWNAVVQSQLT AAPTSQVQ/SDSPTFPSSWDYRHVPEYPANFL *RQGFPMLPRLVSNSWAQTVHPPRPPKVLDL QA
361	1711	A		56	1449	PVPAPRVSPSARGAPGRPRLPGVRGPRHS/WA AD*RGSRM/PPRAPAPSPTGP/APGGKKVRGR VPEDPDAYEPRCSAL*V*PTHVTSPQFCDP*N GQIRSYFTVLLRGLNETMLVK/PLCRREP/PEA GPGRQSTPAVTRDHRQHEDPRGAGRQWDAD PRPSAP/PAEVATGSRPGRHMWMRLCLAAQQ APGLPHRTSIRPGWRRLTEPEAWARRHRRPW GQRGAVRPPPQGAAPPPSHQGRRTNTDPSAT PRLTVMSRCLAPDLKAPASGPRGWRRGMPQ SS/GALLWTPPPTPRGSHSPRPREAPLRAHPA GPSK/SRAGASGRLPEVIYGWVTLFTPPEAGT F/LIPSPT*MSPALVIQPPVPPTQMGLRISGLPR QG*PSGAPW*LPGLAQLAFQCHLPHDEVGPP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	uence	·	09/496	1	to last amino	
seq-	delice			correspondi	acid residue	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first		Q=Glutamine, R=Arginine, S=Serine,
1	1 '		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1			ŀ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1		peptide		/=possible nucleotide deletion, \=possible
1	,		J	sequence	]	nucleotide insertion
	1	i				RNQSPLGNDTLSSGLPMGPRRQVWPLARVG
			1	[	i	GHSSPREPQVLKKPLWGQTDIAGVGSASLYP
Į.	1	l			i	DNL
362	1712	A	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ
302	1732	1	3130	1270	2/4	HNTWQLSRVYPSDLRTDSSNYNPOELWNAG
				ļ	}	
				1		CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR
1		1				VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS
1			İ			LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL
		ļ	l			EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF
1	1					HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE
i	ì	ĺ	ľ.		ĺ	GSIVDPLVKVQIFGVRLDTARQETNYVENNG
	l .	l				FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW
		i				KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG
1.		1	[			ISLRPASIFVYICIQEGLEGDES
363	1713	<del>  c</del>	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA
303	****	١	3137	00	240	PAMNESPLAPHLHQHLVFSVFQVLTILIGV**
364	1714	A	3140	57	418	
304	1714	<b> </b> ^	3140	37	418	SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK
						NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV
1		ĺ				GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI
						SGASLCPVLSPPR*PLQPSSL
365	1715	Α	3145	122	413	LLPYPSLFVFLRQCHFVT\RLECNGVVSAHCN
		}	1			LHLPGSSDSPASAS*VAGTTGVCHHTRLIF\VF
1	[		ŧ.			LV*TGFHYVAQAGLELLTA*S\PPQLPKVVGL
1	1					QA
366	1716	Α	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN
1			5150		~	FIKIKKLCIEGYY/NREPQNGRKIFANYVS\DK
1	i i					
367	1717	A	3152	3	2267	GLMATIYEELLKLSNKLIQ
307	1717	A.	3132	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT
	1					PIEKSDFAAATHPRAFYLSKPDETPNAWMSD
			]			SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS
1 1						TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP
1 1			l i			NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN
1 1	1					VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE
						SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS
						ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR
						QNLKEKHARHIADLRAYYESEINSLKOKLEA
]				ļ		KEISGVEDWKITNQILVDRCGQLDSALHEATS
1						RVRTLENKNNLLEIEVNDLRERFSAASSASKI
1 1						
1 !		- 1				LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF
		ļ				ENAYKLSDDKEAQLKQENKMFQDLLGEYES
						LGKEHRRVKDALNTTENKLLDAYTQISDLKR
1 !		i				MISKLEAQVKQVEHENMLSLRHNSRIHVRPS
			]			RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT
[	• [	ſ	ĺ		ſ	QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST
] ]						SLLIKKQRETSDTPIMRALKELDEGKIFKNWG
	ļ	ļ	ı j		J	TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK
		ļ	ĺ		i	CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR
j		ļ			1	EIMLTPVTVAYSPKRSPKENLSPGFSHLLSKN
		1			l	ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV
j i				1		KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN
{	ľ	ĺ	[	i	İ	DFEYTAKIRTLAETERFFDELTKEKDQIEAAL
i			- 1	ł		SRMPSPGGRITLQTRLNQVKCLSLNLL
368	1718	A	3163	2 -	2350	
500	1/10	^	2102	-	ال ل	EFKSGGCGAGLVAAGAVLVLYPASRAGERT
1				į.	į	RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL
		ĺ	1		ŀ	LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ
1	i		- 1			MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT
1 1	l	- 1	-1	ļ		
		ł	1	}	1	TVPQTQGQTTAQKVSKKTGPRCSTALATGLK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
					÷	QLQDQLRDAQQQVKALGTERTTLEGHLAKV QAQAEQGQQELKNLRACVLELEERLSTQEGL VQELQKKQVELQEERRGLMSQLEEKERRLQT SEAALSSSQAEVASLRQETVAQAALLTEREER LHGLEMERRRLHNQLQELKGNIRVFCRVRPV LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF TMEGGPGGDPQLEGLIPRALRHLFSVAQELSG QGWTYSFVASYVEIYNETVRDLLATGTRKGQ GGECEIRRAGPGSEELTVTNARYVPVSCEKEV DALLHLARQNRAVARTAQNERSSRSHSVFQL QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL ALGPGERERLRETQAINSSLSTLGLVIMALSN KESHVPYRNSKLTYLLQNSLGGSAKMLMFV NISPLEENVSESLNSLRFASKVEPSVLFGTAQS NRKWKTDPDLCVCVCVCVCVCVCVCVCVP MSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP GYIRKIE*1*IYKKN*INFGGVGLLNIVKISILS/K IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD *HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH RVHVPLGAAATVPVHRSHFPR
371	1721	A	3173	770	510	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP ALASQSAGITGVSHLARPQNLYF
372	1722	Α	3180	381	76	RVLHHDNVPAHSSPOKREISQEFOLEIRHLP*S PDLAPSGCFLFLNLKNIFK/GTHFSLVDNVKK TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL QHY*AYVEK
373	1723	A	3181	410		RREVAGPEGKGLLLASAHTMLTPPLLLLPLL SALVAAAIDAPKTCSPKQFACRDQITCISKGW RCDGERDCPDGSDEAPEICPQSKAQRCQPNE HNCLGTELCVPMSRLCNGVQDCMDGSDEGP HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS SFQLQADGKTCKDFDECSVYGTCSQLCTNTD GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP VLLIANSQNILATYLSGAQVSTITPTSTRQTTA MDFSYANETVCWVHVGDSAAQTQLKCARM PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP KGIALDPAMGKVFFTDYGQIPKVERCDMDG QNRTKLVDSKIVFPHGITLDLVSRLVYWADA YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF ENYLYATNSDNANAQQKTSVIRVNRFNSTEY QVVTRVDKGGALHIYHQRRQPRVRSHACEN DQYGKPGGCSDICLLANSHKARTCRCRSGFS LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI YFADTTSYLIGRQKIDGTERETILKDGIHNVE GVAVDWMGDNLYWTDDGPKKTISVARLEK AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW TDWEEDPKDSRRGRLERAWMDGSHRDIFVT SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW TEYRSGSVYRLERGVGGAPPTVTLLRSERPPI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	]	J	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
Į.				peptide	sequence	
1	1	ĺ	<b>l</b> !	• •		/=possible nucleotide deletion, \=possible
<b></b>	<b>↓</b>	<u> </u>		sequence		nucleotide insertion
Į.	1	ļ				FEIRIMYDAQHQQVGSNKCRVNNAGCSSLCL
	1	i	i I			ATPGSRQCACAEDQVLDADGVTCLANPSYVP
	}	·	1			PPQCQPGEFACANSRCIQERWKCDGDNDCLD
	}	l				NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
1	1	}				LCDGDNDCGNSEDESNATCSARTCPPNQFSC
		1	ļ i			ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
	1		1			CFPLTQFTCNNGRCINENWRCDNDNDCGDNS
	1		1			DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
						NDCGDYSDETHANCTNQATRPPGGCHTDEF
		1				QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
						GVTHVCDPSVKFGCKDSARCISKAWVCDGD
1						NDCEDNSDEENCESLACRPSHPCANNTSVC
	1			İ		LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
		1				GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
						CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
1	1		1		]	WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
						KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
	1	•	Ì			VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
	J		J			LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
1						TLLAGDIEHPRAIALDPRDGILFWTDWDASLP
1	1					RIEAASMSGAGRRTVHRETGSGGWPNGLTV
	i					DYLEKRILWIDARSDAIYSARYDGSGHMEVL
[	ĺ		ĺ			RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
1				ū.		KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
1	1		)			QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
1	1					ACPHLMKLHKDNTTCYEFKKFLLYAROMEIR
-						GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
						QRVYWSDVRTQAIKRAFINGTGVETVVSADL
1						PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
	1					RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
	1					
l						WTDGDNISMANMDGSNRTLLFSGQKGPVGL
	1			1		AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
	1		i			AMRSQLGKATALAIMGDKLWWADQVSEKM
1	1		1	ł	l	GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
1						QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
				l		MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
1				ļ	ļ	GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
1.				.	l	TIYWVDMGLSTISRAKRDQTWREDVVTNGIG
1				l		RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
1				ł	· .	SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
						GQYPRIERSRLDGTERVVLVNVSISWPNGISV
				ļ		DYQDGKLYWCDARTDKIERIDLETGENREVV
1	1					LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
1 .	j i				ļ	RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
				;	ļ	OKGTNYCAVANGGCOOLCLYRGRGORACA
	]			ļ	ļ	CAHGMLAEDGASCREYAGYLLYSERTILKSI
1						HLSDERNLNAPVOPFEDPEHMKNVIALAFDY
1						RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT
1						IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
		- 1		ľ		
				-		RHTVDQTRPGAFERETVITMSGDDHPRAFVL
]	]					DECQNLMFWTNWNEQHPSIMRAALSGANVL
	[ [	1	İ			TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
	1		j			RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
1						WTDWVRRAVQRANKHVGSNMKLLRVDIPQ
1		- 1			(	QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
		İ	l	1		LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR
			İ	İ		AQDEFECANGECINFSLTCDGVPHCKDKSDE
1		- 1	ł	i	•	KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
I			>	1	, (	GADDCGDGSDEIPCNKTACGVGEFRCRDGTC
1		1	ļ			IGNSSRCNQFVDCEDASDEMNCSATDCSSYF
L	L					TOTAL TOTAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						RLGVKGVLFQPCERTSLCYAPSWVCDGAND
1 1						CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
1 1						MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
1 1						CQNHRCISKQWLCDGSDDCGDGSDEAAHCE
!						GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
i l						DGADESIAAGCLYNSTCDDREFMCQNRQCIP
1 1						KHFVCDHDRDCADGSDESPECEYPTCGPSEF
						RCANGRCLSSRQWECDGENDCHDQSDEAPK
1 1						NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
1						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
1						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
						TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
						HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
1 1				ļ		TLLKQGLNNAVALDFDYREQMIYWTDVTTQ
						GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
}						DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
						VSSGLREPRALVVDVQNGYLYWTDWGDHSL
						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
1 . 1					1	RIYWADAREDYIEFASLDGSNRHVVLSQDIPH
] ]						IFALTLFEDYVYWTDWETKSINRAHKTTGTN
!!!						KTLLISTLHRPMDLHVFHALRQPDVPNHPCK
1 1						VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
1						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE
						DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
] ]				•		PAFICDGDNDCQDNSDEANCDIHVCLPSQFK
						CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
1 1						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
!						KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
1 1						KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
						PCSESEFSCANGRCIAGRWKCDGDHDCADGS
				•		DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA
		•				DADCMDGSDEEACGTGVRTCPLDEFQCNNT
				i*		LCKPLAWKCDGEDDCGDNSDENPEECARFV
l 1						CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
1 1						GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
						SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
						NASICGDÉARCVRTEKAAYCACRSGFHTVPG
						QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
					j	ARNEMETHNTCKAEGSEYQVLYIADDNEIRS
						LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA
			·			GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
1						RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY
			<u> </u>			WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
						AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
						LRETLVQDNIQWPTGLAVDYHNERLYWADA
	l			l	ĺ	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
	ļ			l	j	FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG
	Ì					LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
	ļ					CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
	}					PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
1	Í			İ	ĺ	QPRYTGDKCELDQCWEHCRNGGTCAASPSG
	ļ	1	1	1		MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
1				1		VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
						NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
						CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
1	i	l	1	j	1	CLTCVGHCSNGGSCTMNSKMMPECQCPPHM
		I				TGPRCEEHVFSQQQPGHIASILIPLLLLLLVL
	I	I	I		I	
1 · 1						VAGVVFWYKRRVQGAKGFQHQRMTNGAM NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DPDKPTNFTNPVYATLYMGGHGSRHSLASTD
374	1724	A		191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL NSQTTTADQLLEFFKQVGEVKFVRMAGDET QPTRFAFVEFADQNSVPRALAFNGVMFGDRP LKINHSNNAIVKPPEMTPQAAAKELEEVMKR VREAQSFISAAIEPGWLHSTSLCNDFLGCF*RR RMYRE*APCTICGTFHLCLIINWDL*LF*AYTA K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSK SHSRRKRSQSKHRSRSHNRSRSQKDRRRSK SPHKKRSKSRERKKSRSSHSRDKRKDTREKI KEKERVKEKDREKEREKEREKEKERGKN KDRDKEREKDREKDKEKDREREREKEHEKD RDKEKEKQDKEKEREKDRSKEIDEKRKKDK KSRTPPRSYNASRRSRSSSRERRRRSRSSSRS PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD HISERRERERSTSMRKSSNDRDGKEKLEKNST S
375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG F\FPNLKKSLRGTH\FSSVKK\TTLT\UNSQDP \WF/\FFYP\SPDLQI\PSS\FR\\GL\ND\\Y\\\H\S\QKC \PDLD\GA\Y\KK
376	1726	Α	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI K
377	1727	A	3201	274	1285	KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM GYKLQDLTDVQIMARLQEESLRQDYASTSAS VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL SSLSTLRDGNWRDGCY
378	1728	A	3202	112	1789	VPGVTESRPSVLRGDHLFALLSSETHQEDPIT YKGFVHKV\ELDRVKLSFSMSLLSRFVGWG* PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN WDAKKGEYVFPAKKKLQEYRVLITTLITAGR LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ FITKLLRNYRSHPTILDIPNQLYYEGELQACA DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD
379	1729	A	3206	432	130	SNLRVWDGIRKPACLTNTSCHS PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK *LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
eotide	seq- uence		09/496		corresponding to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		914	correspondi ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence	]	ļ	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Ì	1	[		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	Sequence	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
				3343333		XKQSNNKGTRASSYXEPDAXEQWKFPHKKL
İ		1				QLPGXTHE
380	1730	A	3207	187	507	GGTGHPHPARPPLSGVGGCQCSHSKPWTAGS
						PEQRDHPAPHKQIEAGQGLPGPQAWGG*KGP
1						AXLLPGPGGGPGPVASLEARAQASSGVTPNG
l						GGRTYPYPTFSSGE
381	1731	Α	3225	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/
						EMQLITSLGLQEFDIARNVLELIYAQTLVWIGI
		1				FFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPS
						KAWRASQMMTFFIFLLFFPSFTGVLCTLAITI
	1		1			WRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRP
						GYLWVVWIYRNLIGSVHFFFILTLIVLIITYLY
1	l					WQITEGRKIMIRLLHEQIINEGKDKMFLIEKLI
1			1			KLQDMEKKANPSSLVLERREVEQQGFLHLGE
700	1000		2020	0.5.		HDGSLDLRSRRSVQEGNPRA
382	1732	A	3238	256	38	LLMIKVSSTCFSCHLHHHHHHHHHHHQGHNS
ļ		l				LFFSLKSSSNSSTLPVYLSYNIILVFSKCLVFDF
202	1722	_	2241	1542	343	LFSNACL KGAPSFVRLYOYPNFAGPHAALANKSFFKAD
383	1733	Α	3241	1542	343	KVTMLWNKKATAVLVIASTDVDKTGASYYG
1						EQTLHYIATNGESAVVQLPKNGPIYDVVWNS
1		l				SSTEFCAVYGFMPAKATIFNLKCDPVFDFGTG
			İ			PRNAAYYSPHGHILVLAGFGNLILQI*AD/IMK
						VWNVKNYKLISKPVASDSTYFAWCPDGEHIL
						TATCAPRLRVNNGYKIWHYTGSILHKYDVPS
1		1				NAELWQVSWQPFLDGIFPAKTITYQAVPSEVP
		i				NEEPKVATAYRPPALRNKPITNSKLHEEEPPO
						NMKPQSGNDKPLSKTALKNQRKHEAKKAAK
)	1	l				QEARSDKSPDLAPTPAPQSTPRNTVSQSISGDP
						EIDKKIKNLKKKLKAIEQLKEQAATGKQLEK
1		i				NQLEKIQKETALLQELEDLELGI
384	1734	A	3242	3	678	IRSPAARSPGLETPTCLLFVIAAIAAVFVDSAIP
			i			RLTQHRPQDGSFPYTILDPPLYLPGQCAPPQP
						LSQCARRVHGEKLRRPTFGPRHRGAGTAKMS
ļ	,		]			ASLVRATVRAVSKRKLQPTRAALTLTPSAVN
ŀ					,	KIKQLLKDKPEHVGVKVGVRTRGCNGLSYTL
		I				EYTKTKGDSDEEVIQDGVRVFIEKKAQLTLL
0.5						GTEMDYVEDKLSSEFVFNNPNIKGTCGCGES
205	1005	<u> </u>	2040	2100		FNI
385	1735	A	3243	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPL
						KEEEILPEPGSETPTVASEALAELLHGALLRR
]		l			· ·	GPEMGYLPGPPLGPEGGEEETTTTIITTTTVTT
		1	l			TVTSPVLCNNNISEGEGYVESPDLGSPVSRTL GLLDCTYSIHVYPGYGIEIOVOTLNLSOEEELL
				_ 1		VLAGGGSPGLAPRLLANSSMLGEGQVLRSPT
	ļ					NRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFP
						PRPAHGDVSVTDLHPGGTATFHCDSGYOLOG
		l				EETLICLNGTRPSWNGETPSCMASCGGTIHNA
		l				TLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRL
	1				}	HLHFERVSLDEDNDRLMVRSGGSPLSPVIYDS
1.						DMDDVPERGLISDAQSLYVELLSETPANPLLL
						SLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPG
						ALATFSCLPGYALEPPGPPNAIECVDPTEPHW
	0.	-				NDTEPACKAMCGGELSEPAGVVLSPDWPQS
	~		.			YSPGQDCVWGVHVQEEKRILLQVEILNVREG
				i		DMLTLFDGDGPSARVLAQLRGPQPRRRLLSS
						GPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPR
		•	- : -			NDTCPELPPPEWGWRTASHGDLIRGTVLTYQ
<u> </u>						CEPGYELLGSDILTCQWDLSWSAAPPACQKI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  MTCADPGEIANGHRTASDAGFPVGSHVQYRC LPGYSLEGAAMLTCYSRDTGTPKWSDRVPKC ALKYEPCLNPGVPENGYQTLYKHHYQAGESL RFFCYEGFELIGEVTITCVPGHPSQWTSQPPLC KVTQTTDPSRQLEGGNLALAILLPLGLVIVLG SGVYIYYTKLQGKSLFGFSGSHSYSPITVESDF SNPLYEAGDTREYEVSI
386	1736	A	3250	5725	3984	GTSTVTMATKKHFSIILNLLGMLLKKDNQDT RKLLMTWALEVAVVMKKSETYAPLFCLPSF HKFCKGLLADTLVEDVNICLQACSSLHALSSS LPDDLLQRCVDVCRVQLVHRGTCIRQAFGKL LKSIPLGVFLSNNNHTEIQEISLALRSHMSKAP SNIFHPQDFSD/VISFILYGNSHRTGKDNWLE RLFYSCQRLDKRDQSTIPRNLLKTDAVLWQW AIWEAAQFTVLSKLRTPLGRAQDTFQTIEGIIR SLAGHTLNPDQDVSQWTTADNDEGHGNNQL RLVLLLQYLENLEKLMYNAYEGCANALTSPP KVIRTFLYTNRQTCQDWLTRIRLSIMRVGLLA GQPAVTVRHGFDLLTEMKTTSLSQGNELEVSI MMVVEALCELHCPEAIQGIAVWSSSIVGKHL LWINSVAQQAEGRFEKASVEYQEHLCAMTG VDCCISSFDKSVLTLASAGCKSASLKHCLNGE SRKSVLSKPTDSSPEVINYLGNKACECYISTA DWAAVQEWQNAIHDLKKSTSSTSLNLKADF NYIKSLSSFESGKFVECTEQLELLPGENINLLA GGSKEKIDMKKLLRNM
387	1737	A	3255	380	76	MDIFLYNCKYQVQTEI*NSIQHIMA\SKKLSRF LKYVHNL*AENYKTLMK*INEDLNKQRDVPY S*TARLNKMSIPTKTIFRFKAIYIKIPATYFIET NMQ
388	1738	A	3260	685	428	PQWLGLQVYALPPANFVFFVEMRSTILAQTG FELLDSSDLPASASKSAGITCMSHHARTLSLK *WPFCLSATQEKFC*PASEGVAW
389	1739	A	3269	1	332	LDGYHTPIYMLNRIIRLPAAL*IISDQTGHALTI LTRLETQMINADYQNKLTLDYLLTTDREVYE PFNLTNYCLHIHNQRLGAYDLG*V*Q/KLAHV PVQV*HGFDPEAMFR
390	1740	A	3270	2	372	GRCHDQNKGKS\DGPDAQAEACGGESTYQEL LVNQNPIGQPLACRRLTRKIYEGIKKAVKPNH SPRGVKKVHKFVNKGEKGIMVLAGDTLGIGV YCLLPCMC*DRKLTYAHIPSTTDLGAGAGY
391	1741	A	3273	1		FFQEMLDIMKAISDMMGKCTYPVLKEDAPRQ HVETFFQELTRSQEGMKLGENFLMFAMPP DDSKESKGK*FFQEMLDIMKAISDMMGKCTY PVLKEDAPRQHVETFFQVGINQKSRGHEVRR KFPDVCHAPR
392	1742	A	3281	901	521	FFFGDGVSPCRQAGV*WHDLDSLQNLPPGFK RFSYLSLPSSW\DYRHVLPRQANFCIF\M*RRG FTMLARMVSIS*PRDLPALASQSAGITGVSHH APPQMDFTFALLCFALKGCLPRQKEGGTLNLI
393	1743	A	3283	385	3	RNRSVVPEFVLLGLSAGPQTQTLLFVLFVVIC LLTVMGNLLLLVVINADSCLHTPMYFFLGQL SFLDLCHSSVTAPKLLENLLSEKKTISVEGCM A*VFFVFATGGTESSLLAVMAYDRYVAIRTR G
394	1744	A	3284	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKC LDNCPEGLEANNHTMECVSIVHCEVSEWNP WSPCTKKGKTCGFKRGTETRVREIIQHPSAKG NLCPPTNETRKCTVQRKKCQKGERGKKGRE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		Ī		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	l		peptide		/=possible nucleotide deletion, \=possible
	l	1		sequence		nucleotide insertion
		i		<u> </u>	1	RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
						KQQQ
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
		1			1	WICLSMVILTHSLKTFHRNWDWESEYTLFMS
	1				İ	ALKVNKNNAKLWNNVGHALENEKNFERAL
L	<b>}</b>					KYFLQATHVQPDDIGAHMNVGR
396	1746	A	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
						IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	A	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
		I				WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
		ļ				AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
	l	l	1			GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
						GLT
398	1748	Α	3300	1912	2768	KQRRWQNIQRKGPKRYTVIAGNSQSHQPMIFS
		!				MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
		1			1	STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
		l				LKGLYGNRDLTARLELFTGRFKDWMVSMIV
		1				DREYSVAVEAVRLLILILKNMEGVLMDVDCE
		l		-		SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
			İ			RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
,		}	j .			SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
			'			VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
399	1749	A	3301	536	2391	VSNS LRSYGCKAPSRISHLHKIFLFLLLPSLLMGYSE
J77	1/47	^	2201	J30	اوري ا	SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
į i		ĺ				CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
		1				GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
		1	}			STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ
		1				SLCGLIFTL/RTPCRPSILHPNYKIISTSAWOKV
						LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
						SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
					!	N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
1						TGAALAGSYPIWENENTLSWLPTFTYNFCLST
		l				PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
		l				NILPPNQTILISVEASISSSPIRNKWALHLITLLT
						GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
						DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
\						TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
						DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
						LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
		ļ.				ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
						P
400	1750	A	3303	2	453	THWRHSSGVPGSTTARRRRELEIATSDNQE
				*		YYNRLCQEVTNRERNDQKMLADLDDLNRTK
						KYLEERLIELLRDKDALWQKSDALEFQQKLS
						AEERWLGDTEANHCLDCKREFSWMVRRHHC
أحجا						RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
				54		CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
			]			EKIVLRALKDSRAGMPEQDKDPGVQENPDD
					*	QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
'						GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
						NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
400	7-02			1600		LSLSTQILG*QKPSKYIPSLCKR
402	1752	A	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
]						PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
						VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
						HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
		поа		, -		
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	)	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	Ì	1		peptide	1	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
	<del> </del>	<del> </del>	1			GAQLKAPLSPLAFRMEDTEALPLTPILYPTCO
ŀ	]			1		FFFFIFLNIFLLAFSSPGSOPLLNSPPSFVCWSR
į	1				ļ	GFMEMNGRGELVESLKRFCASTRLPPTPLLLF
	i	Į		İ		
	•	1			ŀ	PEEEATNGREGLLRFSSWPFSIQDVVQPLTLQ
Ĭ	ĺ					VQRTLVSVTVSDASWVSELL\WSLFVPFTVY
	1		ľ			QVRWLRPVHRQLGEANEEFALRVQQ\LVAKE
1	}	}	ļ	1		LG\QTGTRLTPA\DKAEHMKRQRHPR\LRPQS
!	1	1				AQSSFPPSPWVLSS/SDVQTGQTLGFREFKESF
ł	1					CPHVAIGVFIPERPWPKTGCCKTLTIHLILL*G
	1		1			GPVSFSCPE\DIHPRGT*VPTQQASGLPSFPSYG
	l			I		PARGGVL*HPSAQQPLTFA\KSS\WARAGRAL
		1		1	<b>[</b>	QERKQ\ALYEYARRFTERRAPGGLD
403	1753	Α	3307	44	447	DPSPSLLAVALGLRAGERTRSGPGSSSPSGGIS
		1		l	1	GGASAGLASSPECACGRSHFTCAVSALGECT
			•		1	CIPAOWOCDGDNDCGDHSDEDGCILPTCSPL
1	İ	ĺ	i	İ	į	DFHCDNGKCIRRSWVCDSDNDCEDDSDEQD
	1	1		1		CPPRECEED CPPRECEED
404	1554		2211	400		
404	1754	A	3311	409	1	PRHGWGRRVLGRDRPRLQKVKKSVKAIYIPG
	ì			•		QDHVQNEEIYARVLDKFGSNFLSRDNADLGT
	1				Ì	AFVKFSTLTK*LSALLKNLLQGLSRNVIFTLDS
	l	j	}	] .		LLKGDLKGVKGDLKKPFDKAWKDYETKFAK
						IEKEKREREWR
405	1755	Α	3322	12	458	AAVPVENPWDDPRVRPRVRIFTWEDCIAGQA
	İ			1		KVLCNDSYGVTIDWSPKGAFIRLTSQSVGNG
	l					HPASKENDOMVDTIKNTTKVPIIWTYGDMVE
[	1	ſ	ſ	ĺ		PRPQMIRPAVGAKHKELWKILMALKKIK\IWE
1	İ	ļ	l			GKYTKPSQYNPNYMLELAHNDSVW
406	1756	A	3324	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAV
1 700	1,,50	^	3324	<del>*</del>	720	MCVLLWALSLLQSILEWMFCSFLFSDVDSDN
1	ĺ	ľ	ĺ		1	WCQILDFLTAVWLIFLILVLCGFTLVLLVRIIC
	i				1	
]	ļ		! .		ļ .	GSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*F
					42.72	LLYWIEKDLDDL
407	1757	A	3328	213	1841	SGDLSPAELMMLTIGDVIKQLIEAHEQGKDID
	İ		ŀ	•	ŀ	LNKVKTKTAAKYGLSAQPRLVDIIAAVPPQY
J	ļ				ļ	RKVLMPKLKAKPIRTASGIAVVAVMCKPHRC
]		l	]			PHISFTGNICVYCPGGPDSDFEYSTQSYTGYEP
		l				TSMRAIRARYDPFLQTRHRIEQLKQLGHSVD
		1				KVEFIVMGGTFMALPEEYRDYFIRNLHDALS
		l			- 1	GHTSNNIYEAVKYSERSLTKCIGITIETRPDYC
		1				MKRHLSDMLTYGCTRLEIGVQSVYEDVARD
		1				TNRGHTVKAVCESFHLAKDSGFKVVAHMMP
		[				DLPNVGLERDIEOFTEFFENPAFRPDGLKLYP
		1				TLVIRGTGLYELWKSGRYKSYSPSDLVELVA
				1		RILALVPPWTRVYRVQRDIPMPLVSSGVEHG
						NLRELALARMKDLGIQCRDVRTREVGIQEIH
						HKVRPYQVELVRRDYVANGGWETFLSYEDP
						DQDILIGLLRLRKCSEETFRFELGGGVSIVREL
						HVYGSVVPVSSRDPTKFQHQGFGMLLMEEA
]			-			ERIAREEHGSGKIAVISGVGTRNYYRKIGYRL
<u> </u>						QGPYMVKMLK
408	1758	A	3335	3	467	ALASPRAAGIRHELTSTMAAGKNKRLTKGGK
						KGAKKKAV/DNIINIGKTLVTRTQRTKIASDG
						LKGRVFEESLADLOND\TDGYLLRVI*VAFTT
						ERTNOL/REVFNKLIPDSIGKDIEKACOSIYPLH
·						DDFARKVKMLKKPKFELRKLMELHGEGSS
400	1750	Α	2220	<del></del>	1262	
409	1759	A	3338	7	1252	PRWRNSARDEILLSFPQNYYIQWLNGSLIHGL
						WNLASLFSNLCLFVLMPFAFFFLESEGFAGLK
			'	, ,		KGIRARILETLGMLLLLALLILGIVWVASALID
L				<u></u> _	<u></u>	NDAASMESLYDLWEFYLPYLYSCISLMGCLL _
						·

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met. hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutemic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE QIYIITLEEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL
410	1760	A	3339	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNILL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLUP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPAITPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342		2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHG\RKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKFFLKSLNEGKEGPPLSN TTFPHVLPLITLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

Deptide colide seq	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1765   1765   A   3369   431   315   PPWWWGRLSYREMSUFFLTYNYNAILNKTP   PSSPSI.   A   3362   2   2061   PWSPFLSCHWOGGARGPASSAPELATRIC   PSSPSI   PS	1		l	1 -			
1762   1765   A   3369   431   315   1785   A   3369   431   315   1785   A   3369   431   315   1785   A   3369   431   315   1786   A   3369   431   315   1786   A   3369   431   315   1786   A   3382   2   2061   2006   2007   2							
		1					
Peptide   Possible nucleotide deletion, \ \ Possible   nucleotide deletion, \ \ Possible   nucleotide deletion, \ \ \ Possible   nucleotide deletion, \ \ \ Possible   nucleotide deletion, \ \ \ \ Possible   nucleotide deletion, \ \ \ \ \ Possible   nucleotide deletion, \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	1	ł	l	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			İ			sequence	
1763							•
ATVLEPGGKNTPLFINKYOLDI   SHYKEPL TERIKYO					sequence		
1763   A   3361   3   474   PPYRWNSLEGRILGYEQHANDGKDYISRN   PPLRSWTAADMAAQITKKWEAEEFAAQDK   YLEGTCVERLRITHLENGKETLQLTEQSSQPTI   PIVGIVAGLULLQAVVTGAVSAVMCRIKINS   PPLRSWTAADMAAQITKKWEAEEFAAQDK   YLEGTCVERLRITHLENGKETLQLTEQSSQPTI   PIVGIVAGLULLQAVVTGAVSAVMCRIKINS   GHTLPTDRVSYSEAASSDHAQGSDVSLTACK   YLEGTCVERLRITHLENGKETLQLTGASSSDHAQGSDVSLTACK   YLEGTCVERLRITHLENGKETLQLTGASSSDHAQGSDVSLTACK   YLEGTCVERLRITHLENGKETLQLTGASSSDHAQGSDVSLTACK   YLEGTCVERLRITHLENGKETLQCYGASSESDVL   GVHMVDKDTERDIEMKROLRRILRITHLENGKETLQLTGASSBHAQGSDVSLTACK   YLEGTCVERLRITHLENGKETLQCYGASSESDHAQGSDLGASSHACH   YLEGTCVERLRITHLENGKETLQCYGASSESDHADAGT   YLEGTCVERLRITHLENGKETLQCYGASSISSAPAAEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGDLEDLEE   EHVPGQTVSEEATGVHMMQVDPATLA   KQLEDSTITGSHQGWASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGASSISSAPAEEATEKIK   YLEGTCVERLRITHLENGKETLQCYGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA				į			
1763	İ	i I	ĺ	Í :		ĺ	
1764	413	1763	Α	3361	3	474	
1764	15	1,,	· ·	5501	,	1/1	
1764	1						
1764							PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS
1764	·	·		ì			GHFLPTDRVSYSEAASSDHAQGSDVSLTACK
GVEM-VONDOTERDIEMKROLRRIELEIL ST WKKY QEAMKTSLOVPORENDEGSLGKPLCP PELLSTLI-GSVKRRV-CPFSEDH_EEFI-AEHLP EASNOSLLTVAHADAGTQTINGDLEDLEEHGP GQTVSEEATEVHMMG0PDFATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EDLEEH-POGTVSEEATGVHMMQVDPATLAKSDL EDLEEH-POGTVSEEATGVHMMQVDPATLAKSDL EDLEEH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL EH-POGTVSEEATGVHMMQVDPATLAKSDL FINIT FOR A 3369 431 315 IPWSWVGRLSVRKSILF*LTYNYNAILNKTP PSFSPSL 416 1766 A 3373 42 651 ROPKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSLWAGGAGABGCRGAGAGGABGABGABGABGABGABGABGABGABGABGABGA							
WKKYQEAMKTSLGYPQERDEGELGKPLCP PELISTINGEGSVKKRVCFPSEDHLEEFIAEHLP EASNQSLLTVAHADAGTQTNGDLEDLEEHOP GQTVSEEATGVHMMGDPDTLAELIRDVLQ ELSSYNGEEDDFEVKTSLGYPQRGDLEDLE EHYPGGTVSEEATGVHMMQVDPATLA KQLEDSTITGSHQQMSASPSSAPAEEATEKTK VEEEVKTRKPKKKTRKPKKKSRSRWVVKCWD IPNIF  415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILINKTP PSPSSSL 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCTLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQREKAA PGQGRPLCSLWAGGAGAGPERGAGEGR9DD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPPNSGCPWSEGRAKQPLSCG PKPQCSLPSCLWAGGAGAGPERGAEGRGPDD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPSICACHGPSGRSPL ASPGKIRGWPPGISRRPGGAFGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPSICACHGPSGRSPL ASPGKIRGWPPGISRRPGGAFGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPGGAPAEVRLHNVPHNL RRFPLP*ARGK*PPGGAFGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPGGAPGAGRGRPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRFPLP*ARGK*PPGGAPGAGRGRPGD QAPDPKSGPWGFUFGGAFGRGPSD QAPDPKSGPWGFWSGRGAARABAAPEGAAFGAPA ASAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	414	1764	A	3363	1488	453	
PELLSETLPGSVKKRVCFPSEDHLEFIAEHLP EASNOSLITVAHADAGTONGOLEDLEEHGP GQTVSEBATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEEEUPEEVKTSLGVPCRGDLEDLE EHHYPGQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEEHVPEQTVSEBATGVHMMQVDPATLAKSDL VEEEVKTRKPKKKTRKPKKKSRWNVLKCWD IPNIF  415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTVNYNAILNKTP PSSPSSL APGLGSPLHPILOPQQGASTGLKERKKQNMKG NGNVTLTPLLPAVQCGCCHQPAGRSPLPSSHS APGLCSPLHPILOPQQGASTGLKERKKQNMKG NGNVTLTPLLPAVQCGCCHQPAGRSPLPSSHS APGLCSPLHPILOPQQGASTGLAGERGPSDD QAPDPKSGPWLFPPGLGAPAEVRLHNYPHIL RRPPLP*ARG*PNGSPUSGAGAAGAGAGGAGGAGGAGGAAPAPGPPRPPGGAPAFABAAA PGQGRPLCSLWAGGAGAGAGAAAAPGAGAAGAAAAAAAAGAAGAAGAAGAA			1				
EASNQSLITVARIADAGTQTINGBLEBLEHOP GQTYSEBATEVHMAGDPDTLABLIRDVI.Q ELSSYNGEEDPEEVKTSLGVPQRGDLEDLE EHYPGQTVSEBATGVHMAQVDPATLA KQLEDSTITIGSHQQMSASPSSAPAEBATEKTK VEEDVETRYRKKKTRYPKKKSRSKWNVLKCWD IPNIF  1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQBEKAA PGGGRPLCSHWAGGAGAHQERGAGERGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQNLSCG PKRQCSLBVAGGAGAHQERGAGERGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PNNSGCPWSEGRAKQNLSCG PKRQCSLPSQVPGDTH ASPCKIRGWPPGISKRPGLGGRSFPPGAPRT WRPBARGPSVQSLPPIESSPGAQTTAR*RPGA KNAGRCGGARGPRISLGPPGPPPAPALPAR ASAGAAAALAVGGVRGAGGARGGGYS GHCSGRPTGRTGPGPQGPPMPARPR*ASS TRGSRRGPGSRPARAAPRAGDRGRRVFKV HLRQHTAV*PRLGDATAPPGGAAGGAPAP RIGGGWDCALLPSPFSVSRISESPAPRIGLPALINSY PLKGPGLPPPWQPRTQTGHVITVVPPSGCEH SKSLDAPPWGPSSVSLSLSEPSAPRIGLEBLRVRT VGCGGLASGISSASVSGLSPNRAGGPGQDW EMTYPSWGTQESGGGGSTGPSTGR*GATLPMPS Y*ALGPSANKSEKPLLST\TGR*CCRISLQLA KGIGQLSEPLLNVETARVSMWVTYPRK EEPEEDEDEDDDNNEEEPETCYPFGMKVQV RYGRGKNQKMYEASIKDSDVGGGCSLTJU-MY YCGWNVRYDPWRASIKDSDVGGGCSLTJU-MY YCGWNVRYDPWRASIKDSDVGGGCSLTJU-MY YCGWNVRYDPWRASIKDSDVGGGCSLTJU-MY YCGWNVRYDPWRASIKDSDVGGGRAAQPNA RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINPISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINPISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH RKKIRNKLDKEKDKDEKYSPNCKCPPALGPN PPFGTINNISWKWYFKLDLTDAKNDSTHIKH ETTSILNGLQASESSABEDSEGGEGRAQDMDN	]		1			<u> </u>	
GGTYSEBATEVHMMEGDPDTLAELLIRDVLQ ELSSYNGEFEDDPECKYSLGVPQRGDLEDLE EHVPGQTVSEBATGVHMMQVDPATLAKSDL EDLEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEHVPEQTVSEBATGVHMMQVDPATLAKSDL EDLEHVPEQTVSEBATGVHMMQVDPATLAK KQLEDSTITIGSRQQMSASPSSAPAEBATEKTK VEEBVKTRKPKKKTRKPSKSRWNVLKCWD PNIF  415 1765 A 3369 431 315 PPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSFLHPLOPQOEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAWGERGAEGRGPSD QAPDPKSGPWLFPFGLGAPAEVRLHNVPHNL RRPPLF*ARKF*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQDPRAGGPAGGFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCKPERGRLP ASPPGKIRGWPPGISKGCWSEGRAKQPLSCG PKPQCSLPSQVPGDTH WRPEARGPSVQSLPPTFSPQSAQTTAR*RPGAP KNAGRCGGARGPLSLGPPPGPPAPALPAR ASAGAGAAAALAVGGVRGAGGAGGTGGY GHCSGRPTGRTGPOPQOPGPPMARR*ASS TRGSRRGGGSRARAPAGGDHGRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPAPAP RIGPGWDCALLPSPPFSPSRAKGCAEPEIWDP SPRRGTSPVPSVRSLSSEPANPRLGLPALINSY PLKGPGLPPPWGRTGTGTVITVPSGSCIEH SKSLD/ROPWGAPPAPGSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGCDKCAEPEIWDP SPRRGTSPYPSVRSLSSEPANPRLGLPALINSY PLKGPGLPPPWGRTGTGTVITVPSGSCIEH SKSLD/ROPWGAPPWGPSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGCSD HSSCGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNAAGGPGCGDW EMTYPSWGTQGSGGGSFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDDPRG*DSPLPS GTGTPP*TTTTSPFPPOPSGVYPSRATLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMVYTYFKL 418 1768 A 3398 304 2121 EEEEEEEDDDDDDNEEEFECYPPGMKVQV RYGRGKNQKMYRASIKDSDVGGGFVLYLVH YCGWNVRYDEWIKADKNPAKKH RKKIKNLDKEKDKEDEGEDDDDDDNFRAPADGNEEFFECYPPGMKVQV RYGRGKNQKMYRASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKNPADKNNYPKKH RKKIKNLDKEKDKDEEFECYPPGMKVQV RYGRGKNQKMYRASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKNPADKNNYPKKH RKKIKNLDKEKDKDEEFEGEDDDDDDNARABDEGEDEDDDDDDNARABDEGEDEAGAQDMDN							·-
ELSSYNGEEEDPEEVKTSIGSPPORGDLEDLE EHYPOGTVSEEATGVHMMQVDPATLAKSDL EDLEEHYPEGTVSEEATGVHMMQVDPATLAK KQLEDSTITIGSHQMSASPSSAPAEATEKTK VEEEVKTREPKKKTRKPSKKSRWNVLKCWD IPNIF  415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG GNOVTLTPLLPAVQCGGHLQPAGRSFLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAIPGERGAEGRGPSD QAPDPKSGPWLFPPG_GAPAEVRLHNVPHNL RRPPL*ARGK*PPNSGCPWSGEAKQPLSCG KPQCSLSPSQVPGDTH ARGK*PPNSGCPWSGEAKQPLSCG KPQCSLSPSQVPGDTH ASPGKIRGWPPGISKRPGLGAFAEVRLHNVPHNL RRPPL*ARGK*PPNSGCPWSGRAKQALSCG KPQCSLSPSQVPGDTH ASPGKIRGWPPGISKRPGLGAFAEVRLHNVPHNL RRPPL*ARGK*PPNSGCPWSGRAKQALSCG KPQCSLSPSQVPGDTH ASPGKIRGWPPGISKRPGLGAFAEVRLHNVPHNL RRPPL*ARGK*PPNSGCPWSGRAKQALSCG KPQCSLSPQVPGTHT WRPEARGFSVQSLPFPSPGAAQTTAR*RFGAP KNAGRCGAIRGPLSLGPPPOPPAPAALPAR ASAGAGAAAAAALAVGGVRGAGGAGFGGY GHCSGRPTGRTGPQPQPPPMPARPR*AS\S TRGSRRGPGSRAAAAAPRACDHGRPVRV HLRQHTAV*EPRLGDATAPPGGAGFGAAPA RGGPGWCALLPSPGPFSPSAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLINSY PLKGPGLPPPWGPRTYTGHVTIVVPSGSCIEH SKSLD/RGPWGAPPWGFSSGLCSPKLATAGP PQSWGLCQIGRRGLGGPGLKREET/GLL*GC SMMHANTKGFGVYTSVRCTSHPGGGGCSD HSSCGGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQCDW EMYPVSWQTQESGGGGSFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGPPGPBFSPRAVGCAEPEIWDP SPRRGTSPVSVWGLWEDGPYPRG*DSPLPS GTGTEP*TTTSIPFPPOPGGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSTLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSWVYTYFIK  418 1768 A 3398 304 2121 EEEEEEDEDDDDDNREEEFECYPPGMKVQV RYGRGKNQKMYRASIKDSDVEGGEVLYLVH RKKNKLDKEKDKDEKYSPKNCKPPALGPN PFPQTNPISWKWYPRADKNDPKKHH RKKNKLDKEKDKDEKYSPKNCKPPALGPN PFPQTNPISWKWYPKLDLTDAKNSDTAHIKSI ETTSLINGLQASSSSAGBEGEDERGAQDMDN	1						
BHYPGQTVSEEATGVHMMQVDPATLAKSDL			1				
KQLEDSTITIGSHQQMSASPSSAPAEEATEKTK VEEVKTRKPKKKTRPSKKSRWNVLKCWD IFNIF  415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSL  416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAPGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQPERACGPDAGGRFARDAPGNSLRPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKURGWPPGISKDPGLGRSFPPGFAPRT WPPEARGPSVVSLPSPTSSQAQTTAR*PRGAP KNAGRCGGARGPRLSLGPPPGPPPAPALPAR ASAGAAAAALAVGGVRGAGGARGTGGY GHCSGRPTGTTGPPQGPGPPPMPARR*ASS TRGSRRGPGSSPARAAAAFRAGDHGRRPVRV HLRQHTLAV*EPRLGAPTAPGGAAGGAGAPGAPAP R\GPGWDCALLPSPGPPSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPFWGPRTQTGHVITTVOPSGSCIEH SKSLD/RGPWGAPPWGPSSGLGSPKLATAGP PQSWGLCQIGRRGGPGLKRGFTGLL*CC SMDHANRTKGPGVPTSNRCFSHIPGIGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASGLSPKLAGGPQGDW EMYPVSWGTQESGGGGGFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGPRG*DSPLPS GTGTEP*TPTTSIFPTPPSFPQPSGVYPSSRATLLPMPS Y*ALGPSANKSEKPLLST*YRGLCCRISLQLA KGIGQLSEPILLNYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKNYEASIKDSDVGGEVLYLVH YCGWNRYDEWIKADKIVRPALKNSDTAHIKSI EITSILNGLQASSESABDSRQEDERGAQDMDN			1				•
VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD   IFNIF   IFN			ļ				
415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPSIL 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PQGGRPLCSLWAGGAGAPGERGASEGRESD QAPPPKSGPWLFPPGLGAPAEVRLHNVPINL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH ASPPGKIRGWPPGISKRPGLGGRSPPGFAPRT WRPEARGPSVQSLPPFSPQSAQTTAR*RPGAP KNAGRCGGARGPILSLGPPPGPPAALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGRPTGRTQPQQPGPPMPARPR*ASIS TRGSRRGPGSRPARAAAARAGDHGRPVRV HLRQHTAV*PBRGADATAPPGGAAGPGAPAP RQFGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLALINSY PLKGPGLPPPWGRRGLGFOLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMFAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SFKTGR*VGMLQA GAGSLQGTGGDGVWGLWEDQPRG*DSPLPS GTGTEP*TTTISIPFPPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLST!YRGLCCRISLQLA KGGQLSEDILLNVETAFWSMWVTYFRK 418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWKKDDKNDENNYFKIKH RKKINKLDKEKDKDEK SYPKNCKPPALGPN PPFQTINPISWK WYPKLDTDAKNSDTAHIKSI EITSILNGLQASSSSAEDSEQEDERGAQDMDN							
415 1765 A 3369 431 315 IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP PSFSPS1 416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKONMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQCGASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAPGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPINL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQDPRACGPDAGGRFARDAPGNSLRPPS SPPJGWPGGULLLPR*VPGSELRCGKPERGLP ASPPGKIRGWPPGISKR*PGLGGRSFPPGFAPRT WRPEARGPS*VQSLPPTFSPGSAQTTAR*PGAP KNAGRCGGANGPBLSLGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR*PTGRTGPGPQGPPPMPARPR*ASS TRGSSR*PGSR*PARAAAPRAGDHGRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSSVPSSVSSLSSEPANPRLGLPALLNSY PLKGPGLPPWGPRTGTGHVITVQPSGSCIEH SKSLD/RGPWGPAPPGPSSSGLCSPKLATAGP PQSWGLCQIGRRGLGGFGLKRGET/GLI-GC SMDHANRTKGPGVPTSNRCFSHIPGIGGGCSD HSSCEGHPDLHAGREMPAAPGLSELER*RFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQGSPKTGR*VGMLQA GAGSLQGTGTGDVWGLWEDGPRG*D*SPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPILSTLYRGLCCRISLQLA KGIGGJLSEPLLNVETAFWSMWVTYFKK AKIGNKLDKEKNOKDEK*SPENCKPPALGPN PPFQTINPISWK WYPKLDLTDAKNSDTAHKSI HT68 A 3398 304 2121 EEEEEEDEDDDDDNREEEFECYPPGMKVQV RYGRRGNOKMTEASIKODVEGGEVLJVH YCGWNVRYDEWKADKNVPKIKH RKKIKNKLDKEKNOKDEK*SPENCKPPALGPN PPFQTINPISWK WYPKLDLTDAKNSDTAHKISI EITSILNGLQASSSSAEDSEQEDERGAQDMDN	}						
416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKQNMKG NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAGPGERGAEGGFSD QAPPPKSGPWLFPPGLGAPAEVRLHNVPINL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH EAQPPRACGPDAGGRFAARDAPGNSLRPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGLP ASPPCKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSTTAR*RPGAP KNAGRCGGARGPRLSLGPPPGPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPOPQGPGPPMPARPR*ASS TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPAPA R\GPGWDCALLPSPGFRSFRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTCTGHVTITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLRGCCS SMDHANRTKGPGVPTSNRCFSHIPGGDCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPOQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TTTSIPFFPQPSGVPPSRATLLPMPS Y*ALGPSANKSEKPILSTLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK AGGIGQLSEIPLLNVETAFWSMWVTYFRK AGGIGQLSEIPLLNVETAFWSMWVTYFRK RKKKINKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYRMADKNVPKIKH RKKKINKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYWRADKNVPKIKH RKKKINKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYRADKNVPKIKH RKKKINKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYRADKNVPKIKH RKKKINKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	115	1566		22.50			
416 1766 A 3373 42 651 RQEKMGLGEIGASGVLRSMLKERKKONMKG NGNYTLTPLLPAVQCGCHLQPAGRSPLPSSIR APGLCSPLHPLPAVQCGCHLQPAGRSPLPSSIR APGLCSPLHPLQPQQEASTCFSGTLQGREKAA PGGCRPLCSL WAGGAGAIPGERGAEGRCPSD QAPDFKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARK*PPPGLGAPAEVRLHNVPHNL RRPPLP*ARK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPPGWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSSPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGARGPRLSLGPPPGPPPAPALPAR ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGRPTGRTGPPQGPGPPPMPARPR*ASIS TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGBDATAPFGGAAGFGAPAP RIGPGWDCALTPSPGPSPSPAWGCASPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPFWGPRTQTGHVITTVQPSGSCIEH SISSLD/RGPWGAPSGCSSKLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SIMDHANRIKGPGVPTSNRCTSHIPGGDGCSD HSSCGHPPLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGGGSFKTGR*VGMLQA GAGSLQGGTGGTGTGTYFTTSIFFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGGGSFKTGR*VGMLQA KGIGQLSEPPLLNVETAFWSMWVTYFRK KGIGCQLSEPPLLNVETAFWSMWVTYFRK KGIGCQLSEPPLLNVETAFWSMWVTYFRK RGGGCSSEFFCYPPGMKVQV RYGRGKNQKMYEASKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKVRPADKNVFKIKH RKKIKNKLDKEKDKDEKYSFKNCKPPALGPN PFFQTNPISWKWVPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEGEDERGAQDMDN	415	1765	A	3369	431	315	
NÖNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS APGLCSPLHPLQPQQGASTCPSGTLQGREKAA PGGQGRPLCSL WAGGAGAUPGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPPGWPGQLLRLLPRVPGSELRCCKFERGRLP ASPPGKINGPWGFISKRPGLGGRSSPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGARGPRLSLGPPPGPPPPAPALPAR ASSAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPOPQGPGPPMPARPR*ASIS TRGSRRGPGRPAPAAAAPRAGDHGRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP RGPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLFALLNSY PLKGPGLPPDWGPRTOTGHVITVOPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRGLGGPGLKRGETJGLL*GC SMMDHANTKGPGVPTSNRCTSHIPGGGGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPRRAGGPGQGDW EMYPVSWQTOESGGQGSFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEF*TFTTSIFFPFPGPSGVYPSRATLLPMPS Y*ALGPSANNSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK 418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKURPADKNYKIKH RKKINKLDKEKDNDEKYSFKNCKPPALGPN PFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEGEDERGAQDMDN	416	1766	A	2272	42	651	
APGLCSPLHPLQPQQEASTCPSGTLQGREKAA PGQGRPLCSLWAGGAGAIPGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNYHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPPIGWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGARGPRLSLGPPPGPPPAPALPAR ASAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTQPQPQPGPPMPARPR*ASIS TRGSRRGPSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPAPAP RGPGWDCALLPSPGRSPRAVCGAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRSTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSGLCSPKLATAGP PQSWGLCG/GRRGLGGPGLKRGET/GL*9C SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQDW EMYPVSWQTQESGGQG/SFKTGR*VGMLQA GAGSLQGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEDEDDDDDNEEEEFCYPFGMKVQV RYGRGKNGKMYTEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKOKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	710	1700	Λ	33/3	42	651	
PGQGRPLCSLWAGGAGAPGERGAEGRGPSD QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL RRRPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPPJGWPGQLRLLPRVPGSELRCCKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPFSPQSAQTTAR*RPGAP KNAGRCGGARGPRLSLGPPPGPPPAPPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPGGPPGPPPMPARPR*ASIS TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP RIGPGWDCALLPSPGRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMMDANRIKGPGVPTSNRCFSHIPGGDGCSD HSSCEGHPDLHAGREMPAAPGLSLERVRFT VGCGGLAGISSASVSGLSPNAGGPGQGDW EMYPVSWQTQESGGQG/SFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDINEEEEFFCYPPGMKVQV RYGRGKNQKMYTASIKGSDVEGGEVL/LVH YCGWNVRYDEWIKADKIVRADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
QAPDPKSGPWLEPPGLGAPAEVRLHNVPHNL RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELR.CGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPITSPQSAQTTAR*RPGAP KNAGRCGGA'RGPRLSLGPPPGPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR*TGRTGPGPQGPGPPMPARPR*ASIS TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTTAV*EPRLGDATAPPGGAAGFCAPAP RIOFF@WDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTOTGHVUTTVQPSGSCEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGIGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRTT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGPRG*DSSLPS GTGTEP*TPTTSIPFFPPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEPLLNVETAFWSMVTYFRK  418 1768 A 3398 304 2121 EEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGRKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDDEKYSPRNCKPPALGPN PPPGVINIJSWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	į į						
PKPQCSLPSQVPGDTH  417 1767 A 3382 2 2061 EAQDPRACGAPAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELR.CGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA/RGPRLSLGPPPGPPPAPALPAR ASAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTQPQQPQPPMPARPR*AS\S TRGSRRGPGSRPARAAAARRAGDHGRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRIKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPRAGGPG\QDW EMYPVSWQTQESGGQG/SFKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDQPRG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKN\QKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKINKLLKEKDKEKYSFKNCKPPALGPN PPFQTNPISWK WYPKLDLTDAKNSDTAHKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							-
417 1767 A 3382 2 2061 EAQDPRACGPDAGGRFAARDAPGNSLRPPPS SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKING WPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK EEEEEDEDDDDNNEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKEKPSKNCKPPALGPN PPFGINPISWK WYPKLDLTDAKNSDTAHKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG
SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSQSAQTTAR*RPGAP KNAGRCGGA/RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPOPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPFPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPQQCDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK 418 1768 A 3398 304 2121 EEEEEEEDEDDDDDNNEEEEFFCYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSFGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLSSPKLATAGP PQSWGLCQIGRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQCDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGPRG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK AGIGQLSEIPLLNVETAFWSMWVTYFRK 418 1768 A 3398 304 2121 EEEEEEDEDDDDDNNEEEEFFCYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFGINPISWK WYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	417	1767	A	3382	2	2061	
WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP KNAGRCGGAIRGPRLSIGPPPGPPPAPALPAR ASAGAGAAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALINSY PLKGPGLPPPWGPRTQTGHVITVOPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGJGQLSEPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKKKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
KNAGRCGGARGPRLSLGPPPQPPPAPALPAR ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQDW EMYPVSWQTQESGGQ\G\SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP\RG*DSPLPS GT\GTEP*TPTTSIPFFQPS\GV\PSRATLLPMPS Y*ALGPSANKSEKPLLSFL\PRGLCCRISLQLA K\GIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV R\YGR\GKNQKM\YEASIKDSDVEG\GEVL\YLVH YCGWNVR\YDEWIKADKIVR\PADK\NVPKIKH RKKKNKLDKEKDKDEK\YSPK\NCKPPAL\GPN PPF\QT\NPIS\WK\YYPKLDLTDAK\NSDTAHIKSI EITSIL\NGLQASESSAEDS\EQEDERGAQDMD\N		:					
ASAGAGAAAALAVGGVRGAGGARGTGGY GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSFRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S TRGSRRGPGSRPARAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGQQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDDNNEEEEFECYPPGMKVQV RYGRKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN						•	
TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							The state of the s
R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGQQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	1						
SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	'						
PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGIGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSFLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN						,	
SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPGGDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	1. 1					l	
HSSCEGHPDLHAGREMPAAPGLSELERVRFT VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
VGCGGLASGISSASVSGLSPNRAGGPGQGDW EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDDDDDNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
EMYPVSWQTQESGGQG/SPKTGR*VGMLQA GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEEDDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA KGIGQLSEIPLLNVETAFWSMWVTYFRK  418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN					ļ		GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS
KGIGQLSEIPLLNVETAFWSMWVTYFRK  1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN					l		
418 1768 A 3398 304 2121 EEEEEEEDEDDDDNNEEEEFECYPPGMKVQV RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN							
RYGRGKNQKMYEASIKDSDVEGGEVLYLVH YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	10-1	1760		2200	204	2121	
YCGWNVRYDEWIKADKIVRPADKNVPKIKH RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN	418	1768	Α.	3398	304	2121	
RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN		·					. 1
PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI EITSILNGLQASESSAEDSEQEDERGAQDMDN				1	]		
EITSILNGLQASESSAEDSEQEDERGAQDMDN							·
			ł	1		•	
				•			NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Ì	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		<u> </u>		sequence		nucleotide insertion
					}	NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
		1				YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
1 .		l			•	SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
		-				SLRTTGFYSGFSEVAEKRIKLLNNSDERLONS
				-		RAKDRKDVWSSIQGOWPKKTLKELFSDSDTE
						AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
i						ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
1. 1		ľ				GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
1						VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE
						LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
1						KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
						QQKEGK
419	1769	Α	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
						NGVVLDTQQDQLENAKMEHTNASFDTFFCE
						TRAGKHVPRALFVDLEPTVIDGIR
420	1770	A	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP
1						PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
						VMGFHHVGQAGLELLTSGDLPALASQSARIT
						GVNHCAQPRGHFH
421	1771	A	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS
						ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
1						LYLVSPLENEPKEMLTLSEYHERVRSQGQQL
	•				·	QQLQAELDKLHKEVSTVRAANSERVAKLVF
1 1					•	QRLNEDFVRKPDYALSSVGASIDLQKTSHDY ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
1 1						GNCWAFEGDQGQVVIQLPGRVQLSDITLQHP
						PPSVEHTGGANSAPRDFAVFFLLSFFTHOGLO
1						VYDETEVSLGKFTFDVEKSEIQTFHLONDPPA
						AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT
					'	SEGAEGSAOGPH
422	1772	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN
						YLVSCSMRSPESGKGEPGTARDYTPMGRPPP
1						PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
						QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
						VDTAGAPASPGPDVCE
423	1773	Α	3420	91	706	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
						RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG
						EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
•						KTHLPGFVEQAEALKAKGVQVVACLSVNDA
						FVTGEWGRAHKAEGKVRLLADPTGAFGKET
1						DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA
424	1774	Ā	3421	4	7688	LNVEPDGTGLTCSLAPNIISQL
424	1//4	A	3421	*	/088	RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
1 1					l	FSEKRYVVQVREDVTPGAPVLRVTASDRDKG SNAVVHYSIMSGNARGQFYLDAQTGALDVV
		•			ļ	SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
}						VTVOVLDINDNAPIFVSTPFOATVLESVPLGY
1						LVLHVQAIDADAGDNARLEYRLAGVGHDFP
1 1	}	.				FTINNGTGWISVAAELDREEVDFYSFGVEAR
			ĺ	ĺ		DHGTPALTASASVSVTALDVNDNNPTFTQPE
		- 1				YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
1						QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
						RQYVLAVTASDGTRQDTAQIVVNVTDANTH
1 1		l	ł	l		RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
						DTGENARITYFMEDSIPQFRIDADTGAVTTQA
		1		ł	1,0	ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI
				1		LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV
					·	LQISATDRDSGLNGRVFYTFQGGDDGDGDFI

SEQ ID SEQ ID Met SEC		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of NO: of hod ID		nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- peptide in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- US		corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence 09/4		to last amino	M=Methionine, N=Asparagine, P=Proline,
uence 914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	peptide		/=possible nucleotide deletion, \=possible
	sequence		nucleotide insertion
			VESTSGIVRTLRRLDRENVAQYVLRAYAVDK
1 1	ļ	ł	GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD
		[	VFVEENSPIGLAVARVTATDPDEGTNAOIMY
			QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE
			YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
	1		LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP
1 1 1		}	DISDSLTYSFERGNELSLVLLNASTGELKLSR
1 1 1			ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV
			TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIO
1 1 1		ł	AVAATLATPPDHVVVFNVQRDTDAPGGHILN
1 1 1			VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
			LLTAISAQRVLPFDDNICLREPCENYMRCVSV
	J		LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF
1 1			TGDYCETEVDLCYSRPCGPHGRCRSREGGYT
			CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT
	1	İ	CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
	-		AHSFITFRGLRQRFHFTLALSFATKERDGLLL
			YNGRFNEKHDFVALEVIQEQVQLTFSAGEST
			TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG
			QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS
	1		VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG
	l l		VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
	İ		ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC
			VNQWDAFSCECPLGFGGKSCAQEMANPOHF
	- 1		LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD
	1		GVLLQAITRGRSTITLQLREGHVMLSVEGTGL
	1		QASSLRLEPGRANDGDWHHAQLALGAIGGP
			GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV
	I		GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN
			SLDPSHGESINVEQGCSLPDPCDSNPCPANSY
			CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC
			EHQSVCTRKPSAPHGYTCECPPNYLGPYCET
1 1 1	i		RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
	ĺ		NKTSGECHCKENHYRPPGSPTCLLCDCYPTG
			SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF
			AEVTTNGCEVNYDSCPRAIEAGIWWPRTRFG
			LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF
	1		NCTSITFSELKGFAERLQRNESGLDSGRSQQL
			ALLLRNATOHTAGYFGSDVKVAYOLATRLL
1 1 1			AHESTQRGFGLSATQDVHFTENLLRVGSALL
	1	j l	DTANKRHWELIQQTEGGTAWLLOHYEAYAS
			ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN
			FAGAKLPRYEALRGEOPPDLETTVILPESVFR
			ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
1 1 1 1	1	1	GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK
] [ ] [			RPINTPVVSISVHDDEELLPRALDKPVTVOFR
]     1			LLETEERTKPICVFWNHSILVSGTGGWSARGC
	1		EVVFRNESHVSCQCNHMTSFAVLMDVSRRE
		ľ	NGEILPLKTLTYVALGVTLAALLLTFFFLTLL
			RILRSNQHGIRRNLTAALGLAQLVFLLGINOA
]			DLPFACTVIAILLHFLYLCTFSWALLEALHLY
] ] ] ]			RALTEVRDVNTGPMRFYYMLGWGVPAFITG
			LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP
	1		
	]	Ī	VAFAVSMSVFLYILAARASCAAQRQGFEKKG
1 1 1		}	PVSGLQPSFAVLLLLSATWLLALLSVNSDTLL
			FHYLFATCNCIQOPFIFLSYVVLSKEVRKALK
• •			LACSRKPSPDPALTTKSTLTSSYNCPSPYADG
		·	RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR
		1	EESALNPG\QGPPGLGGIPGR/LCFLGRFKDQQ
, , i i	I	· i	H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenœ		Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS
İ						TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
1						EERLRENGDALSREGSLGPLPGSSAQPHKGIL KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE
1						GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA
					·	GTVDEDSSGSEFLFFNFLH
425	1775	A	3429	155	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS
						RAASVREAEDAPLQPASIHPVSQGSRGPEGSL
ļ						GSAECLPGDPLGARRATRAHSPVPGPPPSLPA AGTAVKRGLQPG*GA/GATSTPGTGAATGGL
ĺ						CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
İ	ĺ					SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV*
	}					KREFQRGPWAGMVILHRISAADPARAPGPDS
						NLQSALQQPATGCSEPAAVYSPPIGLWGA**P
1						EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS
						QAGFSREDRLEQAKLFCRTLEDILADAPESON
						NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE
				į		EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP
426	1776	A	3431	1662	369	LPLRTDFS AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR
720	1770	Α	3431	1002		SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/
						SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP
						AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE
						KAGPHCSRLALTG\SHDFAINFDPENPECEGK
						RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVF1\EISGVIASYRRCLPQIQLYGPTNVAP
						IDNRVAEPAQREQSTGQATKYSVLLVLTDGV
						VSDMAETRTAIVRASRLPMSIIIVGVGNADFS
						DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR
						DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI
						TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ
·						GISPGAPRPCTLATTPSPSP
427	1777	A	3446 .	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
						GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
						ASRPEASGDCRAGRETAMATLEKLMKAFESL KSFQQQQQQQQQQQQQQQQQQQQQQPPPP
						PPPPPPPOLPOPPPOAOPLLPOPOPPPPPPPPP
			1	1	[	GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
						ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
		1				ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPOK
		l				CRPYLVNLLPCLTRTSKRPEESVOETLAAAVP
						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
				Ì		RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
						LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
] · [		J			]	TLHHTOHODHNVVTGALELLQQLFRTPPPEL
						LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
			l			AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
	1				ļ	ESRSDVSSSALTASVKDEISGELAASSGVSTPG
	1		ł		1	SAGHDIITEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSOVSAVPSDPAMDLNDG
-			l			TOASSPISDSSOTTTEGPDSAVTPSDSSEIVLD
	İ					GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
			l	,		FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
					ļ	VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
	. [	ĺ	[			APLVHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD
<u> </u>						THE THEORY OF THE PARTY OF THE

	P-1			r =		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod .	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
	20100		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence			914			
1 1	i		ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1 !				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 !		1		peptide		/=possible nucleotide deletion, \=possible
1 !				sequence		nucleotide insertion
<del></del>	<del> </del>		<del> </del>			TTEYPEEQYVSDILNYIDHGDPOVRGATAILC
1			1			
1 1						GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
1 , }	ł		•			ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
[ ]						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
1			1			LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
1						LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
						IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
1 1	[		l i			l · · · · · · · · · · · · ·
						KLLMHETOPPSHFSVSTITRIYRGYNLLPSITD
]						VTMENNLSRVLAAVSHELITSTTRALTFGCCE
						ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
1				1		KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
1 1						ILAGNLLAASAPKSLRSSWASEEEANPAATK
1						QEEVWPALGDRALVPMVEQLFSHLLKVINIC
						AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
] ·						GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
1					-	GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
						NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
1 1						ATLQDIGKCVEEILGYLKSCFSREPMMATVC
, I						VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
1 1						QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
1 1						SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
i i						NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ
}						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
1						DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
						KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
			i			LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
1 1			·			CHKENEDKWKRLSRQIADIILPMLAKQQMHI
] 1		. 1				
1						DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF
		1				VTPNTMASVSTVQLWISGILAILRVLISQSTED
! !						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
1 1						EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
1					}	KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
		1				GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
		j	1			ARSMITTHPALVLLWCQILLLVNHTDYRWW
		1				· · · · · · · · · · · · · · · · · · ·
į l				l		AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
1				ı		AKLGMCNREIVRRGALILFCDYVCQNLHDSE
1 1		. 1		1	ŀ	HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
<u> </u>				l		AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
						HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
] ]		ļ			j	ACRRVEMILLAANLOSSMAOLPMEELNRIOEY
1		ł		]		LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
j		ĺ	l	1		PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
, I	l	j		ŀ		
j l			1			SQCWTRSDSALLEGAELVNRIPAEDMNAFM
ایا			}			MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
1 1			l	į	1	AREVTLARVSGTVQQLPAVHHVFQPELPAEP
, 1		į	j	1	1	AAYWSKLNDLFGDAALYOSLPTLARALAOY
1 1	ĺ	i	!	1	1	LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
		]	i	l	1	
1		ļ	l	ŀ	1	SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
, I	l	i	l	l	ĺ	WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
	ŀ	l	ľ	l		EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
		l	}	1		TAACEMVAEMVESLQSVLALGHKRNSGVPA
1		ł	1	1	l	FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
			1	l	l	WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
	.		l	l l	I	
1		}	l	İ	· I	INTLGWTSRTQFEETWATLLGVLVTQPLVME
		- 1	į	ŀ	1	QEESPPEEDTERTQINVLAVQAITSLVLSAMT
		- 1	j	l	. 1	VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
1		- 1	Į	ļ	1	LSIIRGIVEQEIQAMVSKRENLATHHLYQAWD
1	l		ľ	ļ	i	PVPSLSPATTGALISHEKULLOINPERFLGSMG
			ĺ		Ì	PVPSLSPATTGALISHEKLLLQINPERELGSMS YKLGOVSIHSVWLGNSITPLREEEWDEEEEEE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of, peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS
						DLFTERNQFELMYVTLTELRRVHPSEDEILAQ YLVPATCKAAAVLGMDKAVAEPVSRLLESTL RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV KLSVDRVNVHSPHRAMAALGLMLTCMYTG KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS TGQSSMVRDWVMLSLSNFTQRAPVAMATWS LSCFFVSASTSPWVAAILPHVISRMGKLEQVD VNLFCLVATDFYRHQIEEELDRRAFQSVLEV VAAPGSPYHRLLTCLRNVHKVTTC
428	1778	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG LPCVGDAAEYQDCNPQACPVRGAWSCWTS WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL GLHTEEALCATQACPEGWS
429	1779	A	3464	583	3	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI PPES/RS*QGGTVQTGQHSSGREAGSWRARGR NAGRR*KGGGKJGTKQGAVRARKECRGEMA SGETDSE
430		A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE ASKCNIVCTQPRRISAVSLANRVCDELGCENG PGGRNSLCGYQIRMESRACESTRLYCTTGV LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS DFLLIILKEILQKRSDLHLIMSATVDSEKFST YFTHCPILRISGRSYPVEVFHLEDIIEETGFVLE KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI LSTQDQAAAFTLPPPGVRKYLATNIAETGITI PDVVFVIDTGRTKENKYHESSQMSSLVETFVS KASALQRQGRAGRVRDGFCFRMYTRERFEG FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF LSKALDPPQLQVISNAMNLLRKIGACELNEPK LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP VATLAAVMTEKSPFTTPIGRKDEADLAKSAL AMADSDHLTIYNAYLGWKKARQEGGYRSEI TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA GLYDNVGKIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK MSLENDKILQIITELIKTENN
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL PATLGGDGGKPALTAGEAALPGLHRSGVPAA AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ QRGEASTGGASGRRCGSCFQV

SEO ID	SEQ ID	Met	T SEQ	Predicted	Predicted end	[A=iid(A-d)-i
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ì			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
			1	peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
432	1782	A	3478	416	23	QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
1				'''	~	QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
	ł	ì		ł	i	CSLFNKWC/YKNWM/LHVOKKRI*VOTLHPS
	ł	1	l	Ì		QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY
						SRALNSGSR
433	1783	A	3504	1876	552	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS
		1				SLGAPGESTLLVRTSKLLVGLGLOLLVWLLL
1	ļ	1	ļ	ļ		QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS
i			1			APRSRCVARPAARTGLPTPAPASSPAPAASPA
			1			PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
			1			GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
		1				AASPSPAASPAPPAASPVLTASPPLPAASPSPA
1	l	1	1			ASPAPPAASPVLTASPPLPAASPALAASPVHT .
		1			1	ASPPVHVASPPVHTASPPVHVASPPVHTASPP
			ļ		1	VHVASPPVHTASPHVHVASPPVHVASPPVHV
ļ		ł				ASPPVHTASPPVHVASPPVHTASPHVHVASPP
						VHTASPPVHVASPPVHVASPPVHVAYPPVHV
İ						ASPPVHVASPPVHVASPPVSCSGDSTSDCFPP
						QPGAVFPHSLAPSLGGWSHLVAALP
434	1784	Α	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
	1	[	{			SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ
						SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ
						WHNLGSLQPLSLEDRLSPGVLGCSALCRSGV
		<u> </u>				RTKFGINMVTSRERGTTRLPKEG
435	1785	Α.	3529	1	3161	MSLVRAALEALDELDLFGVKGGPQSVIHVLA
		l				DEVQHCQSILNSLLPRASTSKEVDASLLSVVS
1		i				FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF
		l				LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC
						EWPLFWTYFILDGVFSGNAEQVQEYKEALEA
	•	[				VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT
	l					VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
		ŀ				PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
	[	ĺ	[ [			SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL
						KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK
-			ļ			
						LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE MLRTDLSYLCSRWRMTGQPTTTFPISHSMLDE
						DGTSLNSSILAALRKMQDGYFGGARVQTGKL
			1			SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN
						YDYLESGNWMNDYDSTSHARCGDEVARYL
			) l			DHLLAHTAPHPKLAPTSOKGGLDRFOAAVOT
						TCDLMSLVTKAKELHVQNVHMYLPTKLFQA
			[ [		I	SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDO
					ļ	SGEVDFKALVLQLKETSSLQEQADILYMLYT
						MKGPDWNTELYNERSATVRELLTELYGKVG
						EIRHWGLIRYISGILRKKVEALDEACTDLLSH
						OKHLTVGLPPEPREKTISAPLPYEALTOLIDEA
						SEGDMSISILTQEIMVYLAMYMRTQPGLFAE
						MFRLRIGLIIQVMATELAHSLRCSAEEATEGL
}		١ .		ļ		MNLSPSAMKNLLHHILSGKEFGVERSVRPTD
		•		ļ		SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK
						QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
						QRRRRLDGALNRVPVGFYQKVWKVLQKCH
				İ		GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL
	' I		<i> </i>	ľ		NRVPQPEYRQLLVEAIL\VLTMLADIEI\HSIGS
					-	IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD
						PASGICTLLYDSAPSGRFGTMTYLSKAAATY
					_*	VQEFLPHSICAMQ
436	1786	Α	3546	73	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL
						EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i i	09/496	согтеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	(	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	]	j		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
						PXSARSCWMRKG
437	1787	Α	3554	5157	2939	AVRAEPGLEELSSGLRAHSPSATTVCEPEAQG
						SASGCRYAAHPHWGLGGAAAAGGSWEPQPP
	ļ	l				RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA
1	ł	l	1			AQPGARARTSPPPASARNMAARPAATLAWSL
		l		:		LLLSSALLREGCRARFVAERDSEDDGEEPVVF
	1	l				PESPLQSPTVLVAVLARNAAHTLPHFLGCLER
1			i i		·	LDYPKSRMAIWAATDHNVDNTTEIFREWLK
1		i		,		NVQRLYHYVEWRPMDEPESYPDEIGPKHWP
		1				TSRFAHVMKLRQAALRTAREKWSDYILFIDV
	•	1				DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS
1		ĺ	<b>i</b> 1			NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP
		l		•	·	VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW
1	ļ	١.				TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP
1						LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGOGGD
1						
1		}				RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA     LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV
			1			WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK
						LMDNIDOAOLDWELIYIGRKRMOVKEPEKA
İ						VPNVANLVEADYSYWTLGYVISLEGAQKLV
1.		ŀ				GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
						YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE
						TSTIWDNETVATDWDRTHAWKSRKOSRIYSN
						AKNTEALPPPTSLDTVPSRDEL
438	1788	A	3563	130	527	IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C
1	2.00					NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA
						NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA
1						SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT
						TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLLSTLTGRSY
					-	GQPSLQDELKDNTTVFTRILDRLLDGYDNRL
				-		RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI
			İ			DVFFRQSWKDERLKFKGPMTVLRLNNLMAS
1						KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE
-		,		1		DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH
						ACPLKFGSYAYTRAEVVYEWTREPARSVVV
			0			AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV
				i		MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF
						WLNRESVPARTVFGVTTVLTMTTLSISARNSL
				'	1	PKVAYATAMDWFIAVCYAFVFSALIEFATVN
						YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN
1		١.,				NTYAPTATSYTPNLARGDPGLATIAKSATIEP
	•					KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
140						FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	Α	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK
						FWEVISDEHGIDPTGTYHGDSDLQLERINVYY
		19				NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG
	,					GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS
						EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ
'					i	FIEELITKWQKNDQELISDPLQQCFKKDEILDG
[ i			ľ	I	ĺ	QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE
					l	VKTPFDLAKAQENSNSVKKKTKFVNLYTREG
j l						QDRLAVLLPGRHPCDCLGQKHKLINNCLICG
	[		ĺ	l		RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS
						N/KSQKLLKKLMSGVENSGKVDISTKDLLPH
						QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						DDESDYFASDSNQWLSKLERETLQKREEELR ELRHASRLSKKVTIDFAGRKILEEENSLAEYH SRLDETIQAIANGTLNQPLTKLDRSSEEPLGVL VNPNMYQSPPQWVDHTGAASQKKAFRSSGF GLEFNSFQHQLRIQDQEFQEGFDGGWCLSVH QPWASLLVRGIKRVEGRSWYTPHRGRLWIAA
						TAKKPSPQEVSELQATYRLLRGKDVEFPNDY PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDS PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH QGAKKGLMKQNKAV
442	1792	Α .	3576	1	2019	MPRSHTGERLCEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMRLSSLTRHMRSHT AIRANEKPYKCKEC\GRAFSLSQILSK\HERSH TGEKPYKCKQCGKTFIYHQPFQRHERTHIGEK PYECKQCGKALSCSSSLRVHERIHTGEKPYEC KQCGKAFSCSSIRVHERTHTGEKPYACK\EC GKAFIS\TTSVLTHMITHNGDRPYKCKECGKA
						FIFPSFLRVHERIHTGEKPYKCKQCGKAFRWS TSIQIHERIHTGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYFRIHERT HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG EKPYECKECAKTFISLENFRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKQ
						CGKAFSCSSYIRIHKRTHTGEKIPYECKECGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFRIRHTRIHNYEKPLEC*QICGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
						DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY
443	1793	A	3578	287	114	DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF
444	1794	A	3582	3335		HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETQFYTAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK
445	1795	<b>A</b>	3584		6169	RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA FPVAFLETHLDKHNIYSIYNTKSSRERAALSLP TNVEDVCPNIPSLEKLMEEIVELAESGIRYTQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isolaucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l ,	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
(				peptide		/=possible nucleotide deletion, \=possible
ļ				sequence		nucleotide insertion
		1				MPHVMEVILPMLCSYMSRWWEHGPENNPER
		ļ				AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE
i				'		GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM
[ ]		ĺ			!	EKLKKKAATVVSEEDHLKAEARGDMSEAEL
1.		1 1				LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
1						QNFVVQNEINNMSFLITDTKSKMSKAAVSDQ
1		1				ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
1						NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS
1		1				NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
1						DPEKTVERVLDIANVLFHLEOKSKRVGRRHY
1		1				CLVEHPORSKKAVWHKLLSKORKRAVVACF
		ł				RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH
						YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHO
]						LILLFSRTALTEKCKLEEDFLYMAYADIMAKS
						CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ
		ĺĺ				ARLHDRGAAEMVLQTISASKGETGPMVAAT
		1				LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF
		i l				FQSLAGLMQSCSVLDLNAFERQNKAEGLGM
						VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH
j						NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVQ
}						ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA
1 1				1		KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
						VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
						KDMVVMLLSMLEGNVVNGTIGKQMVDMLV
1						ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD
		l j				PDGKGVIFKRDFHKAMESHKHYTQSETEFLL
1 1			1	i		SCAETDENETLDYEEFVKRFHEPAKDIGFNVA
1 1			ľ	'		VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP
1 1		i	- }			FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ VKESKRQFIFDVVNEGGEKEKMELFVNFCED
				İ		TIFEMQLAAQISESDLNERSANKEESEKERPEE
1 1			j	j		QGPRMAFFSILTVRSALFALRYNILTLMRMLS
						LKSLKKOMKKVKKMTVKDMVTAFFSSYWSI
	[		1			FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA
		١. ا				KKIKVAELLANMPDPTQDEVRGDGEEGERKP
1	ł		1		}	LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
1 1		1	1	1		EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE
j	ļ			1		KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
		1				KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
	1		1			FWKKIIAYQQKLLNYFARNFYNMRMLALFV
1 1			1			AFAINFILLFYKVSTSSVVEGKELPTRSSSENA
	1	.	l			KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL
1.		ì		1		PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
]	j	J	J	j		LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP
	ļ	1				NNYWDKFVKRKVMDKYGEFYGRDRISELLG
1	ĺ	l	1	İ	ľ	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
		,				VKYQMWKLGVVFTDNSFLYLAWYMTMSVL
	ļ	j	ł	į	ļ	GHY/NNFFFAAHLLDIAMGFKTLRTILSSVTH
	- 1	}	i	ļ	]	NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF
	j	j	ļ			YNKSEDGDTPDMKCDDMLTCYMFHMYVGV
	- 1	- 1		!	1	RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFVI
	1		}			VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
1	1	ł	ļ		ŀ	FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
1		Į	-	1		FLMYLINKDETEHTGQESYVWKMYQERCWE
446	1706	<u> </u>	2502	<del>,                                    </del>	755	FFPAGDCFRKQYEDQLN
440	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
]	ı	1	i		- 1	DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL
		- 1	ł	1	1	SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML
<u> </u>						PRLVSNSWTQAILLPRPPKMLGLQV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion
447	1797	A	3598	1202	1070	LFVGGGPICPEGASGFAPGPAPAPRVGVDAEV GR*V*GAAASQGA/GSLRPRPTGPGHPGAWL QVWGAAAVCAGPAM*/AVRAKRGPRAG*EP NSPWRSGVLAA\RAVGAGPWP*P*PGCS*ARG PSSRSAPGLASGPAAPLLQGVHSSAGPLLCYI NGTLALGLKP**AWGWGEWRPKG
448	1798		3604	3115		FRRKGGGGPKDFGAGLKYNSRHEKVNGLEE GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG LLLVLLGIGFLVWHLQYRDVRVQKVFNGYM RITNENFVDAYENSNSTEFVSLASKVKDALKL LYSGVPFLGPYHKESAVTAFSEGSVIAYYWSE FSIPQHLVEEAERVMAEERVVMLPPRARSLKS FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR GVELMRFTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVTVYNTVL SPMEPHALVQLCGTYPPSYNLTFHSSVQNVL LITLITNTERRHPGYFEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHVKVRFKFFYLLEPGVPAGTCPKDYVEING EKYCGERSQFVVTSNSNKITVRFHSDGSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPVAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV
449	1799	A	3618	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE
450	1800		3620	1		MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VRQEKRMSKATEVMMQYVENLKRTYEKDH AELMEFKKLANQNSSRSCGPSEDGVLRTARS
				i		MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE NGKTNGDPDCEASAPALTLSCLEELSQETKA RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL QVMYPKLCQHWQVIWMMAAVMLVLTVVL GLYNSYNSCAEQADGPLGRSTCSAAQKDSW WSSGLQHEQPTEQ
451	1801	A	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI RHQRIHTGEKPYECKVCGKAFRVSSQLKQHQ RIHTGERPYQCKELKGRGAEMLAVLAVKEQ NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV QCTALGVWTAPAPVCIAVQCQHLEALNEGT MG*DYPFTAFAYGSSCKYECHTVYRVRGLD MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP VCIAVQCQHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWRKGSLLRYRT DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT DNGPVIPVVYDFGDAQKTASYYSPYGQREFT AGFVQFRVFNNERAANALCAGMRVTGCNTE HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT VHVGYSSSREITE\AAVLLFYR
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK F
457	1807	A	3660	14	1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF TKCRSPERETFSCHWTDEVHIGTKNLGPIQLF YTRRNTQEWTQEWKECPDYVSAGENSCYFN SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ PDPPIALNWTLLNVSLTGHADIQVRWEAPRN ADIQKGWMVLEYELQYKEVNETKWKMMDP ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIIF GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH EKLHINLGVKDGDSGRTSCCEPDILETDFNAH DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGQKNKAGMSQCDMHPEMVSLCQENF LMDNAYFCEADAKKCIPVAPHIKVESHIQPS LNQEDIYITTESLTNTAAGSPGTGEHVPGSEM

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Gluternine, R=Arginine, S=Serine.
uaice			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		}	•	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1			1	peptide		/=possible nucleotide deletion, \=possible
		ŀ		sequence		nucleotide insertion
						PVPDYTSIHIVQSPQGLILNATALPLPDKEFLS SCGYVSTDQLNKIMP
458	1808	A	3663	154	462	TRAPASGRSGAGLALSANAPDSGGHPGATEG
1						PAGSLAHASGSARGTWRVRGRGSHGWERTV
}					}	GAGGCANPVPALHSCASAPRGTGRVSALGPK
150	1000		3664	000		TGSSPLSSPKG
459	1809	Α	3664	902	135	LGKYNTSMALFDFVLHNSTGEIRYITEDDVIQ
<u> </u>					Į	SQNALGKYNTSMALFESNSFEKTILESPYYVD LNQTLFVQVSLHTSDPNLVVFLDTCRASPTSD
		i			1	FASPTYDLIKSGCSRDETCK\VYPLFGHYGRF
1						QFNAFKFLRSMSSVYLQCKVLICDSSDHOSRC
						WQGCVSRSKRDISSYKWKTDSIIGPIRLKRDR
[ ]						SAINGNSGFQHETHAEETPNQPFNSVHLFSFM
}		ļ				VLALNVVTVATITVRHFVNQRADYQ\YQKLQ
100	1010		2670	0.50	666	NY
460	1810	Α	3670	850	557	LGILMSPQVEAGEI*ALLTPPPGCMQFSPLTL/P K*WVSPGLTP/PPPEVPSVFLVEPGLPHAGOA
						GLDLL\TSGDPPASTSQSARTTDVSHRAOPLAI
						S
461	1811	Α	3671	2472	2099	IGVLAFETGSCSVTRLYCIGHMPHCSLDLAGS\
						TSAFRIAGTTSVHHHPQLTFFFFWIETGSHCV
						VQTGL*LLALSNPPALASQIAGISGMSHRAWP
100	1010	_	2602	204		GLVLYSLEFSLLCASQSLIMLFTCYNE
462	1812	Α	3672	394	110	VKPVNGESKRD*GADTQTCEGEADEQLQT\N
! !						CYYD/STKSFFYISCG*K\RKPTWAENRRLNA KMFGIPLHSNSDPWGYEEREVIGFHRSRVSRG
						HGS
463	1813	A	3673	348	1	QRNPFSAGHPQRPPTSGSQSELLAQPRLRPGR
! !						KSSFSRDQDVW*SQAVPKRQ*QRNPFSAGHP
j j						QRPPTSGSQSELLAQPRLRPGRKSSFSRDQDV
						WPGQKPRPSQQQHQMCASPTLGQRSPFALEP VPAYHGGRDPFASARPSPVGIPKPRAAPAGG
						GWRRIRPKSSTK
464	1814	A	3676	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSR
						KVFQLLPSFPTLTRSKSHESQLGNRIDDVSSM
						RFDLSHGSPQMVRRDIGLSVTHRFSTKSWLS
						QVCHVCQKSMIFGVKCKHCRLKCHNKCTKE
}						APACRISFLPLTRLRRTESVPSDINNPVDRAAE PHFGTLPKALTKKEHPPAMNHLDSSSNPSSTT
						FSTPSSPAPFPTSSNPSSATTPPNPSP\GQR\DSR
						FNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPE
	j					AADGTRLDDQPKADVLEAHEAEAEEPEAGK
					İ	SEAEDDEDEVDDLPSSRRPWRGPISRKASQTS
		1		[	ĺ	VYLQEWDIPFEQVELGEPIGQGRWGRVHRGR
[	ſ	ſ		ĺ	ĺ	WHGEVAIRLLEMDGHNQDHLKLFKKEVMN
		Ì		ļ		YRQTRHENVVLFMGACMNPPHLAIITSFCKG
ì	}		1	l	ľ	RTLHSFVRDPKTSLDINKTRQIAQEIIKGMGY LHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF
	{	.		.!	ŀ	\GISGVVP\EGRRENQLKLSHDWLCYLAPEIVR
		i		l		EMTPGKDEDQLPFSKAADVYAFGTVWYELQ
	1	ļ		ļ		ARDWPLKNQAAEASIWQIGSGEGMKRVLTS
		j		J	-1	VSLGKEVSENLSACWAFDLQERPS\FSLLMD
	j		J			MLEKLPKLNRRLSHPGHF*KSADINSSKVVPR
465	1815	A -	3679	8	803	FERFGLGVLESSNPKM
403	1013	^	30/7	٥		IPSPAWWNSTWADTFSLLLALAVALYLGYY WACVLQTHRAFCASNTEDLETVVNHIKHRYP
		- 1	Ì		,	QAPLLAVGISFGGILVLNHLAQARQAAGLVA
	[	Í	[			ALTLSACWDSFETTRSLETPLNSLLFNQPLTA
						GLCQLVERLSY/E*DLQARTIRQFDERYTSVA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide seq-		USSN	nucleotide location	location .	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
1 2020			,	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		FGYODCVTYYKAASPRTKIDAIRIPVLYLSAA
1						DDPFSTVCALPKQAAQHSPYVALLITARGGHI
						GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
	1016		2604		207	GLPDLRALLPSEDRNS
466	1816	A	3684	3	307	SSQYTVQSKTKIFL*AAREKQ/RHTCRRFSIRLS ANISSQTGEARGQWPSVFKVLKEKKLSTKKS
						FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
						TGVLQG
467	1817	Α	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA
						RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF
						FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE
						APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG
						MARARLAQLVRLAGGHCRRDTLWKRLFLLE
						PPGPDRLRLGGRLALAELEELLEAVHAKSIGD
						IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR HFOSPDLGTOYLVVLNOKFTDCFVLVFLDSH
						LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY
	,					HHLESVINTACFTLWTRLL*GSGLDH*MSLFL
						ESWAYQIACQRQD*PALLGPRASQTLSDTKG
i		_				FVTMS*GSAAPAWQQEPPSPNTHSH*PIQDSR   ESGOPRGPLGPFWGTPFGPPGRVSGVHTGWO
						TPPRAPLPESCPLYPLTTVSHLCPLSLRVFTSHL
						DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA
					}	LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG
468	1818	A	3691	960	499	GKQQRN OTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI
400	1010	^	3091	300	422	QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH
						DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF
ì						PGNF\*FLVKTGFPHVGQTGFELLTSSDLAPLA
469	1819	A	3714	4747	495	SQNGGITGMSPCAWPFFFFFFGLC MAYSWQTDPNPNESHEKQYEHQEFLFVNQP
409	1019	A	3/14	4/4/	493	HSSSOVSLGFDOIVDEISGKIPHYESEIDENTFF
						VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL
				•		SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG
						SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL
						ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK RSGHVNIVEPSLMLLKGSLQPGMWESTWQK
						NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR
j				1		ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK
						FNIHIFIDNSTOPLHFMPCANYLVKDLIAEILH
						FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS HKMFOKDKSVIQLHLQKSREAPGKLSRKHEE
				,		DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY
.						DFHLKYLLKTQENVYNIJEEVKKICSVLGCVE
						TKQITDAVNELSLILQRKGENFYQSSETSAKG LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV
						PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR
				'		INFPLEIKSLPRESMLTVKLFGIACATNNANLL
				:		AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM
1						ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD
					•	SEENRSNLEEPLKECIKHIARLSQKQTPLLLSE EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
						DERTVSEMHTILRRWTFSQPLEALGLLTSSFP
[·						DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV
}						KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL
				.		LKNAENEAYFKSWYQKLLAALQFCAGKALN DEFSKEOKLIKILGDIGERVKSASDHOROEVL
		0				KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH
						DACSYFTSNALPLKITFINANLMGKNISIIFKA
						<del></del>

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible nucleotide insertion GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ MITYRCLSTGKDQRLVQMVPDAVTLAKIHRH SGLIGPLKENTIKKWFSQHNHLKADYEKALR NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS EMEYFITEGGKNPQHFQDFVELCCRAYNIIR KHSQLLL\NLL\EMMLYAG\LPELSGI\QDLKY VYNNLRPQDTDLEATSHFTKKIKESLECFPVK
470	1820	Α	3718	430	75	LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST TRSIERATILGFSKKSSNLYLIQVTHSNNETSL TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY NEIVYYDEVTELQGHVLMLIVKSKTVFVGAI NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM KSLEKDEFVGGMLLSNPIW SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA
						VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/ CPWGARAEGRLSAVVDSQISCCKGDLCNAV VLAAGSPWALCVQLLLSLGSVFLWALL
471	1821	A	3723	891	494	LRQSL/NSVPQAGVQWRDSSVPLWALL LRQSL/NSVPQAGVQWRDSSLQAPPPRFTPLS CLSLPSSWDYRRLPPCLANFLYF**RRGFTML ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF SKNPI
472	1822	Α	3734	443	251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN *ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473	1823	A	3746	3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF LITLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE PPPPVRQEPSPKPNNKTPAILYTYSGLRNRA AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV KFAENRAK
474	1824		3753	2	. 5262	RPLFAREGGIYAVLVCMQEYKTSVLVQQAG LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ MTREIFASIDSATRPGSESLLLTVPAAVILMLN TEGCSSAARNGLLLNLLCNHHTLGDQIITQ ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR YSEPPGSPERAALETPIIQGQDGSPELLIRSLV GGPSAELLLDLERVLCREGSPGGAVRPLLKRL QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL DQYLEQRETSRNPLSRAASFASRVRRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP VLPSSSLRNITQCWLSVVQEQVSRFLAAAWR APDFVPRYCKLYEHLQRAGSELFGPRAAFML

SEQ ID	'SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	}			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
)				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ
	1	i i				QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG
						LELATTFEHFYQHYMADRLLSFGSSWLEGAV
	·					LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF
						QLQRLDKLFLEQEDEEEKRL*EEEEEEEEEA EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP
Ì	ľ					RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG
1 :		Į i			}	PHRRLOWTWLGRAELQFGKQILHVSTVQMW
		1				LLLKFNQTEEVSVETLLKDSDLSPELLLQALV
!						PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
1						RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
						LSCLLVRILKAHGEKGLHIDOLVCLVLEAWO
1	]	Į į				KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL
						GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
						DVPFCGSQSETSKPSPEAVATLASLQLPAGRT
1						MSPQEVEGLMKQTVRQVQETLNLEPDVAQH
						LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
1	ļ					HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC
						MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD
						CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE
						SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
						CGWASCFNCSFPEAHYPASCGHMSQWVDDG
						GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
			'		,	KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
						KDYYNCSAMVSKAARQEKRFQDYNERCTFH
						HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
1 .						CQGLEQARKVLAYACVYSFYSQDAEYMDVV   EQQTENLELHTNALQILLEETLLRCRDLASSL
		[				RLLRADCLSTGMELLRRIQERLLAILQHSAQD
						FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP
						EAEEEEEDDEDDVPEWQQDEFDEELDNDSFS
						YDESENLDQETFFFGDEEEDEDEAYD
475	1825	Α	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ
'''	1023	**	3,3.	10,5	70	ATGRRRRTRTQQRTAALLTDGTTKTGAAW
						SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN
1						PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP
						DGTRIRPASITGVAQSPATRATPSLPCLHVPAP
						SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
						RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE
						PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS
						HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
	:	1				PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN
						TLWEEGRQRPPETLQPAR
476	1826	Α	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK
						RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
						FTMLARMVSIS*PRDLPALASQSAGITGVSHH
177	1000		27/1	042		APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG
						ARTTPS*IFVFLVETGFHHVSQDGLDLL/NFVI
470	1000		2762	267	1240	RPRRPLKVLGLQACTRARLPSPLKEL
478	1828	Α	3763	267	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
]		<i>'</i>				PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS
						PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
		'				
						WTGHSFASQAWLRQVPEVSKHLQCPSAESLL TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG
						PDEHPQDTDARDADGEAREREP/RRPSFAA*P
		-				VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA
				.		CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE
				.		QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA
L						Z Zor

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EGSDHIG
479	1829	A	3766	2	2152	YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSPGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFPFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRPILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR FCEAMDNYFLMVPDHCTGLGLNC
480	1830	A	3777	251	3	QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNI\KDPAITLDVYPNEVKNYVRTKTYTQMF I/ANFIMAKSWKQPTHPSVRT
481	1831	<b>A</b> _	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	A	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPPRLPKCWDYRREPPRP A*FFVFLVE\QGFTMLARMVSIS*PQ\CDLPAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTPFSCLSLPG SWDYRRPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	A	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSCLSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771	320	FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghttamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	•	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ĺ	ſ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
			i	peptide	0040000	/-possible nucleotide deletion, \-possible
				sequence	ŀ	nucleotide insertion
488	1838	A	3818	1	781	FRACLLELIPYAPTLSWTACPPAMAGPRGLLP
1		l	l	· ·	1	LCLLAFCLAGFSFVRGQVLFKGCDVKTTFVT
			]		:	HVPCTSCAAIKKOTCPSGWLRELPDOITODCR
			Ì			YEVOLGGSMVSMSGCRRKCRKOVVOKACCP
			Ì			GYWGSRCHECPGGAETPCNGHGTCLDGMDR
1	ł					NGTCVCQENFRGSACQECQDPNRFGPDCQSV
1						CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
ł						QELPVWQELGFPQNNPRLRKAPNCKCLPG*H
L						RNGLIATPNPCRP
489	1839	Α	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP
1		ļ .				ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
	[					QDGLDLL/NLMIHPPRPPKVLGFQA
490	1840	Α	3825	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
	1	l				GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
] .		]				ASRPEASGDCRAGRETAMATLEKLMKAFESL
1						KSFQQQQQQQQQQQQQQQQQQQPPPP
						PPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPPPPPP
		1				GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
		i		'		ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
1						ESDVRMVADECLNKVIKALMDSNLPRLQLEL
						YKEIKKNGAPRSLRAALWRFAELAHLVRPQK
						CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
1	,					RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
						LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
į į						KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
1						TLHHTQHQDHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
						AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS
1 1						ESRSDVSSSALTASVKDEISGELAASSGVSTPG
1						SAGHDUTEQPRSQHTLQADSVDLASCDLTSS
1						ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
	2					TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
1						GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
! :						FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
i l						VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
						APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
					_	SVKALALSCVGAAVALHPESFFSKLYKVPLD
						TTEYPEEQYVSDILNYIDHGDPQVRGATAILC
			94		Į	GTLICSILSRSRFHVGDWMGTIRTLTGNTFSL
	1			1	İ	ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
1						SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
]						LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
			[	ſ	[	LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
			1	ļ		IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
						KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
	ĺ		1			VTMENNLSRVIAAVSHELITSTTRALTFGCCE
}					ĺ	ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
	1	ł	1	ľ	ł	KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
	1					ILAGNLLAASAPKSLRSSWASEEEANPAATK
						QEEVWPALGDRALVPMVEQLFSHLLKVINIC
	İ		1			AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK
]	1	1	- 1	1	ļ	GKEKEPGEQASVPLSPKKGSEASAASRQSDTS
				1	Φ.	GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
		İ	1		1	NYKVTLDLQNSTEKFGGFLRSALDVLSQILEL
			- 1		1	ATLQDIGKCVEEILGYLKSCFSREPMMATVC
			j	]		VQQLLKTLFGTNLASQFDGLSSNPSKSQGRA
]	ŀ		İ			QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
1 1		1	ĺ	ı	[	SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
<b></b>		1	1			NLTSVTKNRADKNAIHNHIRLFEPLVIKALKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
nence			914	ng to first	acid residue	Q=Ghrtamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
1	Į į					DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
1	:					KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
		ŀ				LETQKEVVVSMLLRLIQYHQVLEMFILVLQQ
						CHKENEDKWKRLSRQIADIILPMLAKQQMHI
					•	DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF
1						VTPNTMASVSTVQLWISGILAILRVLISQSTED
						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE
						EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
1						KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
						GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
1						ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
						AKLGMCNREIVRRGALILFCDYVCONLHDSE
						HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
1						AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
						HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
						ACRRVEMLLAANLOSSMAQLPMEELNRIQEY
ſ						LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
						PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
]						SQCWTRSDSALLEGAELVNRIPAEDMNAFM
			·			MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
1						AREVTLARVSGTVQQLPAVHHVFQPELPAEP
						AAYWSKLNDLFGDAALYQSLPTLARALAQY
				}		LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
1						SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
1 1			1			WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
	ŀ					EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI
						TAACEMVAEMVESLQSVLALGHKRNSGVPA
						FLTPLLRNIISLARLPLVNSYTRVPPLVWKLG
		1				WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
1				•	,	INTLGWTSRTQFEETWATLLGVLVTQPLVME
		į				QEESPPEEDTERTOINVLAVQAITSLVLSAMT
						VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
						LSIRGIVEQEIQAMVSKRENIATHHLYQAWD
	.					PVPSLSPATTGALISHEKLLLQINPERELGSMS
]	ļ					YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE
] [	1	ſ	ĺ		<i>'</i>	ADAPAPSSPPTSPVNSRKHRAGVDIHSCSOFL
						LELYSRWILPSSSARRTPAILISEVVRSLLVVS
		-	•			DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
	ļ	l	1	1		YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
	1	Ì	l	.		RSSHLPSRVGALHGVLYVLECDLLDDTAKOL
1 1	1	l	1	1		IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA
1		1		i		
		l		l		TAFYLIENYPLDVGPEFSASIIQMCGVMLSGS EESTPSIIYHCALRGLERLLLSEOLSRLDAESL
	}	i				
1	ł	*	ł	i	ł	VKLSVDRVNVHSPHRAMAALGLMLTCMYT GKEKVSPGRTSDPNDA APDSESVIVAMERYS
.		ľ				GKEKVSPGRTSDPNPAAPDSESVIVAMERVS
				ļ		VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ
	1			į	ĺ	DIMNKVIGEFLSNQQPYPQFMATVVYKVFQT
	-				ļ	LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA
	ļ	ļ	J	j	J	TWSLSCFFVSASTSPWVAAILPHVISRMGKLE
			.	ŀ		QVDVNLFCLVATDFYRHQIEEELDRRAFQSV
401	1041	,	2026	460		LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
400	1040	_,	2006	202		HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	88	VAPSPMIMPDLYFYRDPEEIEKEE*AAAEK\EE
( 1	1	ł	1			FQSEWTAVV/P/EFTATQSEVADWFKDMQVP
			. 1			SVPIQQFPTEDWST*PTMNDWSATSTAQTTE
L						WVRITTEWP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł	1	1	1	peptide		/=possible nucleotide deletion, \=possible
1		i		sequence		nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
1	J	ļ				KFKRS*EKAHIRYKIDQPEDIPLK\EFLCKHSK
ŀ	l	ì				CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
	ŀ		ŀ	ł		VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL/ELGLLOAPPPGF
ŀ	1			1	1	KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
	j	ĺ	ļ		i .	FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
						ARPQDIDFLYAHQGRCWFRLL
495	1845	A	3847	1774	40	DIFFRAKEGMGQDEAQFSVEMPLTGKAYL
1		'		1 ****	"	WADKYRPRKPRFFNRVHTGFEWNKYNOTHY
		l	1			DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
		1		[		EACADNKDFAILRFHAGPPYEDIAFKIVNREW
				ŀ	1	EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
	l	1				RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
1	i	İ	ł	ł	ł	
1		Į.				QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC HGELRRHWDRLA*GPDATEGALGASFEHEG
i	l	l	i	1	i	
						GQQPPADLTVQADTLHRPSARLGGAHRACPK
1		l	1		!	RRPHRVLWRWARGAWAWRCQAREKQETQG
1	}	1	İ		i	QPCHITGHPLGREAEPAAAGAAPALAHRPPF
1				}		ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
1	1		}	Ì	1	WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
1	}					VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE
1	j .	ļ	]		J	KGEGLPPGIPPGLLTGPW\SMRPVTPSFAHIR
						TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
	ļ		<b>[</b>			LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
i	t					SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
496	1846	A	3849	830	442	SFVLMELAYWQDRMFF
490	1040	Α.	3049	630	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
			[ [			LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG
			1			LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
497	1847		2050		200	RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
491	1847	A	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
1		1				LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS
		1			·	PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR
				·		RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
400	1040		2062			VPIL
498	1848	A	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG
						EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP
1 1	'	i i		ļ	0	LPCLANF\*FLVETGFHHVGQADLKLLTSGDP
100	1040		20.65			PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	Α	3863	423	263	·APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
						KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
				,	-	KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
						MTFFTEIEKSIIKFTWNHKKPPNTQSNIEQKE*S
			1	j	ĺ	FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
						LVGIALNL
500	1850	Α	3865	2	15246	LPRGCLWCLQRSPTPARPQPSRPARSPLPLFP
				1		DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
		(	i	1	. 1	VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
		1	ŀ			TSNAQNVPPDLAICCFVLEQSLSVRALQEML
			[			ANTVEAGVESSQGGGHRTLLYGHAILLRHAH
	1	l	ì	ł		SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
				l	I	ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
	1	1		1	ł	SERYLHLSTASGELQVDASFMQTLWNMNPIC
	ł	1	1	i		SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
1 1	l	-	}	}	ì	DDQRRLVYYEGGAVCTHARSLWRLEPLRIS
j		į		I		WSGSHLRWGQPLRVRHVTTGQYLALTEDQG
	-	ļ		I	1	LVVVDASKAHTKATSFCFRISKEKLDVAPKR
<u> </u>			}	j		DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	<b>i</b>	'		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ĺ		peptide		/=possible nucleotide deletion, \=possible
<b>}</b>	<del> </del>			sequence		nucleotide insertion  APDPKALRLGVLKKKAMLHQEGHMDDALSL
1						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
1						GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
1						EDLOHEEKOSKLRSLRNROSLFOEEGMLSMV
						LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
]						VNLLYELLASLIRGNRSNCALFSTNLDWLVS
						KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
İ						RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
						IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
						ATHLRVGWALTEGYTPYPGAGEGWGGNGV
						GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
ì						LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
						NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
i		,				LERIREKLAENIHELWALTRIEQGWTYGPVRD
						DNKRLHPCLVDFHSLPEPERNYNLOMSGETL
			1			KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
						NGHNVWARDRVGQGWSYSAVQDIPARRNPR
1						LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
1						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
						QSGRWYFEFEAVTTGEMRVGWARPELRPDV
						ELGADELAYVFNGHRGQRWHLGSEPFGRPW
1					,	QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
						VSSLRFFAICGLQEGFEPFAINMORPVTTWFS
						KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
		1				LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
						CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
					•	NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
		ĺ	Ì			AGGEAQPARAENEKDATTEKNKKRGFLFKA
1					ł	KKVAMMTQPPATPTLPRLPHDVVPADNRDD
		ļ				PEILLNTTTYYYSVRVFAGQEPSCVWAGWVT
		- 1		ĺ		PDYHQHDMSFDLSKVRVVTVTMGDEQGNV HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
·		i	ĺ	İ		VIGCLVDLATGLMTFTANGKESNTFFQVEPN
		- 1				TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
		l				AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
] ]		ļ		ļ		MPNHFLQVETRRAGERLGWAVQCQEPLTMM
			į			ALHIPEENRCMDILELSERLDLQRFHSHTLRL
		1	ļ			YRAVCALGNNRVAHALCSHVDQAQLLHALE
j j	ļ ļ	j	1	j	j	DAHLPGPLRAGYYDLLISIHLESACRSRRSML
			İ	,		SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
	l	ł				ARLSPAIPLEALRDKALRMLGEAVRDGGOHA
			l			RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
	- 1	- 1		ľ		DVKOILKMIEPEVFTEEEEEEDEEEGEEEDEE
]	1	- 1	ì			EKEEDEETAQEKEDEEKEEEAAEGEKEEG
		l			ļ	LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
1	ł	ł	ł	1	I	HRVESLAAFAERYVDKLQANQRSRYGLLIKA
	ļ	1			. [	FSMTAAETARRTREFRSPPQEQINMLLQFKDG
	1	[				TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
		- 1	]	J	Ý.	GEEEPEETTLGSRLMSLLEKVRLVKKKEEK
	·			l	ĺ	PÉEERSAEESKPRSLQELVSHMVVRWAQEDF
					.	VQSPELVRAMFSLLHRQYDGLGELLRALPRA VTISPESVEDTMSLLECT COPPELL INCOME DOE
						YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
[	1	- 1	1	ı		TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL
LL						THE THAT THE COURSE FROM THE TOCKEL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	0		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
				residue of		
1					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						CYFCRISRQNQRSMFDHLSYLLENSGIGLGM
1						QGSTPLDVAAASVIDNNELALALQEQDLEKV
						VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
						ERYLDFLRFAVFVNGESVEENANVVVRLLIR
						KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
1						DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
1 1						FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
						LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
]					0	MSASFVPDHKASMVLFLDRVYGIENQDFLLH
1						
1						VLDVGFLPDMRAAASLDTATFSTTEMALAV
1						NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
1						MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
						YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
		' I			1	TNHYERCWKYYCLPTGWANFGVTSEEELHL
1						TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
						AIAGALPPDYVDASYSSKAEKKATVDAEGNF
					İ	DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
1		•				DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
]						KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
1						EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
1 1				·		SAVTLSRELQAMAEQLAENYHNTWGRKKKO
i I						ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
1						QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
[ ]	-					FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1						EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
1 1	! }					TPAKVLGSGGHASNKEKEMITSLFCKLAALV
	•		i			
						RHRVSLFGTDAPAVVNCLHILARSLDARTVM
						KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
						KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
1 1	}					QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
	1		1	-	l l	TTKNTYVEKLRPALGECLARLAAAMPVAFLE
						PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
i I	.					CPDIPVLERLMADIGGLAESGARYTEMPHVIE
i l			}			ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
l · [						CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
1	ĺ	- 1	i	ĺ	ĺ	KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
1			ļ	J		AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
				J		VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
	]			J		AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
1				J	ļ	NEINNMSFLTADNKSKMAKAGDIOSGGSDOE
		l	1		1	RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
]	. ]	J		ſ		MCAPTDQDLITLAKTRYALKDTDEEVREFLH
	. 1		.1	1		NNLHLQGKVEGSPSLRWQMALYRGVPGREE
1	-		İ	. ]		DADDPEKIVRRVOEVSAVLYYLDOTEHPYKS
	]		J			
			ļ			KKAVWHKLLSKQRRRAVVACFRMTPLYNLP
1 , 1						THRACNMFLESYKAAWILTEDHSFEDRMIDD
	' 1	1	l			LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
		Ì	ļ			ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
			Ì			GENGEAEEEVEVSFEEKQMEKQRLLYQQARL
		*	. 1		1	HTRGAAEMVLQMISACKGETGAMVSSTLKL
	.				i	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
		1	ŀ	ļ	ł	IQALMQTCSVLDLNAFERQNKAEGLGMVNE
!	İ	- 1		ļ		DGTVINRQNGEKVMADDEFTQDLFRFLOLLC
		}		ļ		EGHNNDFQNYLRTQTGNTTTINIIICTVDYLL
	l	ļ				RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
	j	1	ľ	s .	!	SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
j	J	]	}	- 1	ļ	DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
[		- 1	1	-	1	LDLOKDMVVMLLSLLEGNVVNGMIAROMV
		l	[	ľ	İ	7
] ]				ŀ		DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
<u> </u>						QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid residue of peptide	corresponding to last amino acid residue of peptide sequence	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion  QFLLSCSEADENEMINCEEFANRFQEPARDIG FNVAVLLTNLSEHVPHDPRLHNFLELAESILE YFRPYLGRIEIMGASRRIERIYFEISETNRAQW EMPQVKESKRQFIFDVVNEGGEAEKMELFVS
						FCEDTIFEMQIAAQISEPEGEPETDEDEGAGA AEAGAEGAEEGAAGLEGTAATAAAGATARV VAAAGRALRGLSYRSLRRRVRRLRRLTAREA ATAVAALLWAAVTRAGAAGAGAAAGALGL
						LWGSLFGGGLVEGAKKVTVTELLAGMPDPT SDEVHGEQPAGPGGDADGEGASEGAGDAAE GAGDEEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGLGDMGDTTPAEPPTPEGSPILKRKLG
						VDGVEEELPPEPEPEPEPELEPEKADAENGEK EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS
						GGSSGWGLGAGEEAEGDEDENMVYYFLEES TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP LVIFKREKELARKLEFDGLYTTEQPEDDDVKG QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG
						DIYGRERIAELLGMDLATLEITAHNERKPNPP PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV
					:	AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV KEDMETKCFICGIGSDYFDTTPHGFETHTLEE
						HNLANYMFFLMYLINKDETEHTGQESYVWK MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	A	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD YRHAP\PLLTNF\*FLVEMGFCYVGQAGRKLL ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	Α .	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL
504	1854	В	3896	279	70	L MVSKSKSILMSYNHVELTFSDMKKMPEAFRR TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY OKEYYENO*
505	1855	A	3899	2	1396	EPGVPTKKTWFDKPDFNRTNSPGFQKKVQFG NENTKLELRKVPPELNNISKLNEHFSRFGTLV NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS ASSDLPQVLSTYLLA*QKQCIIQLL/WKAAQKT
					,	LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA
L		}			***************************************	VHGRGRGRGRGRGVPGHAVVDHRPRALEIS

SEQ ID NO: of nucl- conde seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA VITFKTRAEAEAAAVHGARFKGODLKLAWN
506	1856	A	3911	1952	919	KPVTNISAVETEEVEPDEEEQREIIIA  DAELSGTLSLVLTQCCKRIKDTVQKLASDHK DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA DSQRLLNEVMVEHFFRQGMLDVAEELCQES GLSVDPSQKEPFVELNRILEALKVRVLRPALE WAVSNREMLIAQNSSLEFKLHRLYFISLLMG GTTNQREALQYAKNFQPFALNHQKDIQVLM GSLVYLRQGIENSPYVHLLDANQWADICDIFT RDACALLGLSVESPLSVSFSAGCVALPALINIK AVIEQRQCTGVWNQKDELPIEVDLG*KSAGY HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL NKMFNGSKLKCPYCPMEQSPGDAKQIFF
507	1857	A	3936	439	18	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA PGSGARCHPPSTCSPSWASPG*GAKASPALPR SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVQ RPAKQRDKRNRHLGR
508	1858	A	3944	120	412	WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT TRHGASRGGVCAHSIEGGELFERIVDEDYHLT EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP MVFFLQNFC/RIILNVA\WTGD*PNTL*KEQRG ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT SASSVAATTG
511	1861	A	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV RWQPSEKQPPPPAHRGPADSLSTAAGAAELS AEGAGKSRGSGEQDWVNRPKTVRDTLLALH QHGHSGPFESKFKKEPALTAVARTARKRKPS PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L LHPLLCLRHHPLPHLIPTGPHRLKRPRMP\SP MAALILVADNAGGSHASKDANQVHSTTRRN SNSPPSPSSMNQRRLGPREVGGQGAGGTGGL EPVHPASLPDSSLATSAPLCCTLCHERLEDTH FVQCPSVPSHKFCFPCSRQSIKQGASGEVYC PSGEKCPLVGSNVPWAFMQGEIATILAGDVK VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR VAGTTDTHHHTWLILGSSVQTGFDHVGQAG LELLTSGDPPISASESAGIMGMSHCVWP*SWG LSHHMAPPQGDGGRARGTPGPEQSFWNLSC H*PRCQVPS*LMTQL/FWGRHQYNPTMKRGK LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPLFHHGL*AWLWCPELLLQGQARRH*RSPPS/FK CPATLSLTAWSQTKRLRSQFLLPWL*RAL*H PP/CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP PSRPDRSRNSNSLSR
513	1863	A .	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTTDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT WDSWASNHSSLHTNHHYNTYHPDHCRVPAL TFPKAPPPNSPPALVSSSFSPTSMSAYSLSSLN MGTLPRSLYSTSPRGTMMRRRLKKKDFKSSL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO; in	beginning nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ	ĺ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		Ì		peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						SLASSTVGLAGQVVHTETTEVVLTADPVTGF GIQLQGSVFATETLSSPPLISYIEADSPAERCG
		l			1	VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI
	İ					TSKVTLEIEFDVAESVIPSSGTFHVKLPKKHN
					İ	VELGITISSPSSRKPGDPLVISDIKKGSVAHRT
		1				GTLELGDKLLAIDNIRLDNCSMEDAVQILQQC
}		ļ			ļ	EDLVKLKIRKDEDNSDEQESSGAIIYTVELKR
					1	YGGPLG\TISGTEEP\FDL*IISSLTKGGLAERT
		İ				GAIHIGDRIL\AINSSSLKGKPLSEAIHLLQMAG
		·				ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD
İ						SWDGSA\IDTS\YGTEGT\SFQASGY\NFNTYD
		1	]	i	]	WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED
}		}	}		1	WDRSTASGFAGAA\DSAETEQEENFWSQALE
						DLETCGQSGILRELEATIMSGSTMSLNHEAPT
{				[		PRSPAGSDRPSFQERSSSRPHYSQTTRSNTLPS
						DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT
						LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA
		'			· ·	GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQOSLPG
						D*SEQNSAFFQQPSHGGNLETREPTNTL
514	1864	A	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG
	100.	••	550.	555	000	NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH
						SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI
						LSEFCMELTGIKQAQVDEGVPLKICLSQFCK
						WIHKIQQQKNIIFATGISEPS/DF*SKIMCICYL
	10/6		2000	400	100	VR*RISYTY*SKHKSKGC
515	1865	Α	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL
			l l			DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIF\*QVKHNPCHRLTPEEGTVQL
}						NRADS
516	1866	A	3977	2	1357	KMLC/OKESNYIRLKRAKMDKSMFVKIKTLGI
						GAFGEVCLARKVDTKALYATKTLRKKDVLL
						RNQVAHVKAERDILAEADNEWVVRLYYSFQ
}						DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL
						ARFYLAELTCAVESVHKMGFIHRDIKPDNILID
						RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP
						RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL
]				- 2	Ì	CDWWSVGVILFEMLVGQPPFLAQTPLETQM
						KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE
						DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS
						AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN
						EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF
1 1						FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD
517	1067		2000	1250	1022	QNTGSEIKNRDLVYV
517	1867	A	3980	1358	1022	FFFKKFTQSLGFLLPSFSFLPSCFFFFHFVLFCY
					İ	VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA
						QAIFQPQPPKVLGLQV
518	1868	A	3986	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F
ļ <sup></sup>						SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH
] ]		J				VGQAGLELLTSGDLPALASQSAGITG\SHRAR
						PENGFENIF
519	1869	A	3994	751	126	NQGLRHVGLCRTCLVNQMFASSILGKSHHHS
, ,		ı	)			LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC
				ļ		DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS
1 1	{		. 1	i		QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP
L						THE SERVICE WALTEN OWALLY I SWAALLST IP

SEO ID	SEQ ID	Met	Tern	Predicted	Dradintal J	Amino said company / A - Aloning C- Outside
NO: of	NO: of	hod	SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
		noa	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ł	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	İ	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i		ľ	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł			į	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ļ		peptide		/=possible nucleotide deletion, \=possible
L		1		sequence		nucleotide insertion .
		·				LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
L		]			J	AHV\FADLLLIITLPSYYIPFC
520	1870	Α	3999	882	698	QSFRLSLLSSWDYRHM*PRLANF*T\FFCRDR/
						SLALLPRLVSNSWPQAILPPRPPKVLGLQT
521	1871	A	4011	1346	1178	FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
}						PPTSASHVAGATGTHHHAWLLSV
522	1872	A	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
322	10/2	1 ^	1013	-	1 377	IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
Į į					i	EYGPVYSTWSALEGELAEPLEGVSACIGNCST
į l	1					
500	1000	<del> </del>	40.0	241		AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
		1				KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
	[	ĺ	1		ĺ	L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
	ĺ					GNEGDITSFPAK
524	1874	Α .	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
	l					TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
		1			,	RQGFTVLARMVLIS*PHDLPASASQSAGITGL
. 1	}	l	J			SHCSWPTSSILS
525	1875	A	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
323	1075	^	7023	701	331	PPRLKO/F/SHLSPPSIWDYRRVPPCLVNFSIFF
	ļ	ļ				
			1			VETGSCQPCLQLLGSSNPPASASQSAGIAGISH
		1	1			QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
						NKIINRPTHPVESSF
526	1876	Α	4024	80	341	TPSSTSRGTEEQQSSKMAWQRREEKEHLNVR
1		l				RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ
						LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	Α	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
						DVAVYFTTKEWAIMG\PAERALYRDVMLEN
						YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
		İ				TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	A	4028	1160	242	GTSELLCIORWNWGPAFPPRPGLALAPTLOLL
		• •		1100	<b>5</b> .5	VEMGSAKSVPVTPARPPPHNKHLARVADPRS
1 1		ł	i i			PSAGILRTPIQVESSPOPGLPAGEOLEGLKHAO
						DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE
		l .				
		l :				VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
						QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
						ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
						NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
						AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
لـــــــــــــــــــــــــــــــــــــ						WEQGQD\HDKENQHFPLVES
529	1879	Α	4039	2	366	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
						CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI,*T
						ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
						HTEICT*MFIAVLFVVVKTWKQF
530	1880	A	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
					-	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
				j		/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
531	1001		4063	-50	070	VKISILPKVNYRFYLISIKIIMAI
121	1881	Α	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY
						HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
		l				IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
						T*KR
532	1882	Α	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
		·				YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
ļ					l	KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL
						QIWMPVSLMNIVTLKCPT
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKU
		1	.5,5	-		ITNI/PFIIASKRIKYSGISLTKEMKDLYTETLLR
						KIKEDTNKWKDI/SCFWVGR/LNIVKMPK/VIC
						THE WALDISCE M A QUALITA VIAIL VALUE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
534	1884	A	4088	3	1931	IFNAIPIKMPMMCMAKIEKNSS  IIDSSTRRMESERSPLYRQLIDLGYLSSSHWNC GAPGQDTKAQSMLVEQSEKLRHLSTFSHQVL QTRLVDAAKALNLVHCHCLDIFINQAFDMQR DLQITPKRLEYTRKKENELYESLMNIANRKQE EMKDMIVETLNTMKEELLDDATNMEFKDVI VPENGEPVGTREIKCCIRQIQELIISRLNQAVA NKLISSVDYLRESFVGTLERCLQSLEKSQDVS VHITSNYLKQILNAAYHVEVTFHSGSSVTRM LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ LEAGHSGRLEKTEDLWLRVKDHAPRLARLS LESRSLQDVLLHRKPKLGQELGRGQYGVVYL CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA VLLIMERLHRDLYTGLKAGLTLETRLQIALDV VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRRGARPERLPVFDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS\ NSEQPNRGLDDST
535	1885	A	4090	2	417	ALMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HNRKRIWLRA
536	1886	A	4102	569	829	DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV
537	1887	A	4104	54	281	SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE
538	1888	A	4109	141	314	IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS
539	1889	A	4111	268	1	ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL
540	1890	A	4142	198	2064	PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFFARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFILEKGLIVPMENVATIADCASVI EGVSRSRNALLNGDTKNYDWDSGYTCHQLG
541	1891	A	4146	282	778	SGAIVVQLAQPYMIGSIRVLLWDCDDRSY GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, i=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion HAESENFAFWQDMKWKNKFWGKSLEIVPVG
						TVNVSLPRFGDHFEWNKVTSCHNVLSGQRW IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG PTPGGQCIWKP
542	1892	A	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT ANVSSENNTPRTSKTTFQLELSVKDAVYTVV SSH
543	1893	A	4153	678	11	TISYPOCLTOMYFLISFANVDTFLLPIMALDH YVAICSALQ*CSIITP/ELCQGLPVLA*AGSSLIS PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA AAILRIPSPTRRKACSICSSHLSLVTLFYGTV LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE LKVGREGHVLPWQAHVVEF
545	1895	A	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM QDTASAMPCLPYYPTSHCFMAGGKSRSQGW ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFLIFFKTEFHSCCPGAVQWHDLDSL QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA GITGVSHHA*PRASGRRCW
547	1897	Α	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP GSHHASEGLAPGVPPAGGVSAQELTAPPQEG WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS GFHSFPRPHQEPSPRSSCWQHLLWHCPWPQ PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS TMAPIPSALAVWEPAGSSPQLSSAPADSSVPLP ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDQI/ RCPPACSPAAASSFSFESQPCPSAPSKASPAPA ALIVGPHHPP*SQQPQSQSVHPHGPGGPQPPL AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA LASVPLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/ PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSSPGPPLPVPALPS WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL PQLPSLRSGSQQPLLFFQCPGPGAVWGKGSPQ PLSPHPPPP/ARTQTFPVASRSLSPGTAPYSVCL TPSRSASSLPEVVLASSLPKIPQSSGSVPLGPTSP MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, '=possible nucleotide deletion, \=possible nucleotide insertion  GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA PGTWQPW/PVSPAPLSCLSAWDPWELPSPQPQ VCSTAELPTSCLLSSPGPAFQPPRFGCL*GPP GPPGLPPLQSSLSFPPPPPPPVQPPAPPALQWG
548	1898	A	4180	2369	844	LHLPGGRTK RHREEDFOFILKGIARLLSNPLLQTYLPNSTK KIQFHQELLVLFWKLCDFNKVGQPRGALQGD GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR ADQSRVGLMHIGVFILLLLSGECNFGVRLNKP YSIRVPMDIPVFTGTHADLLIVVYHKIITSGHQ RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP THKALQRRRTTPEPLSRTGSQGGAPPWRAPA PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR MAPCGPWNLSPSRAWRMAARLRGSPARHGG SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ HGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE LLVRKWRVKSALGAMGQWQLEVGDPAPLG AGNLGPELIKESNANPIFMRKDTKMSFQWRIR NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK KFSIPDLDRHQLPLDDALLSFA\TPTAP
550	:	Α	4192	1	1980	IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL GASAMRRSEVLAEESIVCLQKALNHLREIWE LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE SLKERLIKSISVCQKELNTLCSELHVEPFQEEG ETTILQLEKDLRTQVELMRKQKKERKQEUKL LQEQDQELCEILCMPHYDIDSASVPSLEELNQ FRQHVTILRETKASRREEF/VSSIKRQIILCME ELDHTPDTSFERDVVCEDEDAFCLSLENIATL QKLLRQLEMQKSQNEAVCEGULRTQIRELW DRLQIPEEEREAVATIMSGSKAKVRKALQUE VDRLEELEKCKTMKKVIEAIRVELVQYWDQC FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR LKNYYEVHKELFEGVQKWEETWRLFLEFER KASDPNRFTNRGGNLLKEKQRAKLQKMLP KLEELKARIELWEQEHSKAFMVNGQKFME YVAEQWEMHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK PVAASTCSGKKTPRTGRHGANKENLELNGSI LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	A	4194	3	1008	AWHEGLVSSPAIGAYLSASSYGDSLVVLVATV VALLDICFILVAVPESLPEKMRPVSWGAQISW KQADPFASLKKVGKDSTVLLVCITVCLSYLPE AGQYSSFFLYLR\QVIGFGTVKIAAFIAMVGI LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML QLAWYGFGSQAWMMWAAGTVAAMSSTTPP AISALVSRNAESDQQGVAQGIITGIRGLCNGL GPALYGFIFYMFHVELTELGPKLNSNNVPLQ GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

NO of mucl- coide coide seq- ucacc    NO of file	SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence cité seq uence (USSN location corresponding purpose uence (USSN location corresponding plat in mo acid residue of peptide sequence (Part of peptide sequence) (Part of peptide se			,				
Sequence   19/4   19/	1						
Seq			ł				
usece     914   ng no first amino acid residue of peptide pe			l				
mimo acid residue of peptide sequence physics with properties of peptide sequence physics of period sequence physics of the period sequence period sequence period sequence period sequence period sequence period sequence period sequence period sequence period sequence period sequence period sequence period seq	, .	u u u	l				
residue of peptide sequence   V=Tyrosine, X=Uhixhown, *=Stop podon, /*possible muelocidi deligni, *poss	ucnce	1		714			
peptide sequence  1902 A 4197 2 14302 AAPPFAPGSRQVKOKAAPGAAAAALERGAR PAPARARGTMANGGGEGEDIGHTENDEVV LQCTATHIKEQQKLCLAAEGGRRLCFLESTS NSKNVPPDLSICTFTLOSISVRALOGGBALATI VEKSEGQVDVEKWKFMMKTAQGGGHRTLL VGHALLBRYSVGMVLCLASEGGEKVQR VGDDLLVSVSSERVLHLSVYMGSLTNDAAF QQTD.WSVAPSSGSEAAQGVLIGGDVLRLH GHMDGLTVPSGGHEGERKTVHVEGGAV VGDLLVSVSSERVLHLSVYMGSLTNDAAF QQTD.WSVAPSSGSEAAQGVLIGGDVLRLH GHMDGLTVPSGGHEGERKTVHVEGGAV VARSLWRLETLRVAWGSHRWGDYFRLA HTGESTLTVPSGGHEGERKTVHVEGGAV VARSLWRLETLRVAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGDYFRLA HTGESTLDAWGSTRAWGSHRWGGSHRWGGSHRWGGAV VARGEAGSWKSILNSLYELAALBRAWGSHRWGGSHRWGGAV VAGREAGSWKSILNSLYELAALBRAWGSHRWGGSHRWGGAV VAGREAGSWKSILNSLYELAALBRAWGSHRWGGSHRWGAV VAGREAGSWKSILNSLYELAALBRAWGSHRWGGAV VAGREAGSWKSILNSLYELAALBRAWGSHRWGGAV VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSILNSLYELAALBRAWGSHRWAD VAGREAGSWKSINSTAWGSHRWAD VAGREAGSWKSINSTAWGSHRWAD VARACAASSAWGSHWAD VARACAASSAWGSHWAD VARACAASSAWGSHWAD VARACAASSAWGSHWAGSWAGDWAGSWAGDWAGDWAGSWAGDWAGDWAGDWAGSWAGDWAGDWAGDWAGSWAGDWAGDWAGGAAALBRAWGSHAGDWAGGAAAACACCS GTGERFRIPAEKTYAACGAWGSHGADWAGGAWGAGDWAGGAAAACACCS GTGERFRIPAEKTYAACGAWGSHGADWAGGAWGAGDWAGGAAAAACACCS GTGERFRIPAEKTYAACGAWGAGAWGGAGAAAAACACCS GTGERFRIPAEKTYAACGAGAWGSHGADWAGAAACACCS GTGERFRIPAEKTYAACGAGAAAACACCS GTGERFRIPAEKTYAACGAGAAAACACCS GTGERFRIPAEKTYAACGAGAAAACACCS GTGERFRIPAEKTYAACGA	1	1	}				
Sequence   Sequence						sequence	
1902 A 4197 2   14302   APPPAPOSROQICKAAPGAAAALERGAR   PAPARRRGTHANGGGEGEDEQHERDEVY LUCTATHIKEGOKLCL.AAEGGRRLCFLESTS NISKNYPPDLSICTY*LLOSISSVALLOPEMANT VEKSEGOVDVEKWKPMMKTAGOGGHRTLL YCHAILLRSYSGMYLCLSTSNSSTNKLAPD VGLOEDTTGEACWTHIPASKQRSEGEKVR VODDLL VSVSSERY*LH.SYGNGSLEMANT VEKSEGOVDVEKWKPMMKTAGOGGHRTLL YCHAILLRSYSGMYLCLSTSNSSTNKLAPD VGLOEDTTGEACWTHIPASKQRSEGEKVR VODDLL VSVSSERY*LH.SYGNGSLEMADOAYLOGDVIRLH GHMDECLTYPSGEHEGEGRETYHTPGAAPS VYHARSI WRLETLR VAWSGSHRWOOPPRLR HYTTGKY*SLAMEDNILLADIKKAAVSTA FTFRSSKKLDVGVRKEVDGAGTSEKVODS VCYQHVDTOLWLTYSSAMEDSOR KAIMHEGHMDDGISLSRSQHEESNTAAVIRS TYLE-PARREGILDALSKAKAASTYDLFISVSIS SLQDLIOYFHYPDEHLEHEDKOPRLRALKAR QNLFOSEGMMIN WELDFLALFALDFAND CAGESGMIN WELDFLALFOX SEGGEMUN LUCTOLWLTYSSAMEND CAGESGMIN WELDFLALFOX SEGGEMUN LUCTOLWLTYSSAMEND CAGESGMIN WELDFLALFOX SEGGEMUN LUCTOLWLTYSSAMEN WELDFLAND CAGESGMIN WELDFLAND SEGGEMUN CAGESGMIN WELDFLAND CAGESGMIN WELDFLAND SEGGEMUN CAGESGMIN WELDFLAND CAGESGMIN WELDFLAND SEGGEMUN CAGESGMIN WELDFLAND SEGGEMUN CAGESGMIN WELDFLAND SEGGEMUN CAGESGMIN WELDFLAND SEGGEMUN SEGGEMUN WELDFLAND SEGGEMUN WELDFLAND SEGGEMUN SEGGEMUN WELDFLAND SEGGEMUN W	1		ł				
EPAPARRIGTIÁ-ADGGEGEDEJRI-RTDDEVV LQCTATHIKEQOKLOL-AAFGGRALCHESTS NSKNYPPDLSICTFVLEQSLSVRALQEMLANT VEKSEGOVDYEW-KRY-MMCTAGOGGHTLL YCHAILLRISYSGNYL-CLSTSESSTIKLAFD VOLOEDTTOEA-GWYNTHAP-SIKQRSEGEK-VR VODDLIL VSVSSER YL-HLSYGNGSLHVDAAF QQTIL-WSVABFISGSBA-AQG'TLIGGDV/RLLH GHMDECLTYPSGEBEGEQRATVHYEGGA'S VRAASLWRLETLRYAWGGSHRWGQFFRLR HYTTGRY-LSLMEDEN/GLLMDKEKADVSTA FTRSSKEKLDVOVRE-VDGMGTSEIKY GDS VCYGQUYDTGL-WLT-YGSDVSVKSWGSIGR KAMMHEGHMDDGISLSKSQHEESKTAAVVRS TVLFFRFRIEGLALSKK-KASTVDLFESVSL SLQDLIGYFHPDEHLEHEDKONRLR-ALKNR QNLFQEEGMNI-VLSCHBLH-PUSSAAHFAD VAGRE-AGESWSLINSLYELLAALIRGNRN CAGFGSISD-WLISRLER-BASGILE-VLHCVL VESPEALMIKEGHISSISLLDKHGRNIKVLD VLCSLC-VCHGVAVRSNOHLICDNLLPGRDLL LQTRL-VNHYSSMRPNIFL-GVSEGSAQVKK-WY YELM-VDHTEPPTVAE-ATH-RWGWASTEGYSP YPGGGEEWGGNOVGDDLFSYGPGGLH-WSG GLATVYSSPNQHLLBTDUSTCCLDLSAPSISF RINGQPVQGMEENFINDLLFPTVSFSAGIKV RFLLGGRHGFKK-IPDVAYCFAVLPKK-K KVEHSREYKQRTYTTDLLGFTVSLTQAAFT PPVTTSSYNLPH-LERREKALENHIEL-WVM KIELGWQYGPVRDDNKRQHFCL-VFFSKLPEQ ERNYNLGMSETLKLLALGCHGUSDEHAE DKVKKMKLPRNYOLTSGYR-PAMDLSFIKLT PSGEAM-VDLALBGHGWHY-ROW SPREED HAD KVKKMKLPRNYOLTSGYR-PAMDLSFIKLT PSGEAM-VDLALBGHGWSPL-VFFSKLPEQ ERNYNLGMSETLKLLALGCHGUSDEHAE DKVKKMKLPRNYOLTSGYR-PAMDLSFIKLT PSGEAM-VDLALBAR-BOYS GUGDVKNRRYPRL-VYTPLDDRTKKSNKDS LRA-VETTLAGGTGNSGWAT-ROW SPREED HAD GOOD VKNRRYPRL-VYTPLDDRTKKSNKDS LRA-VETTLAGGTGNSGWAT-ROW SPREED HAD KVKKMKLPRNYOLTSGYR-PAMDLSFIKLT PSGEAM-VDLALBAR-MEYS GTGER-RUFFREKKLAENHIEL-WVM KIELGWQYGP-VRDDNKRQHFCL-VFFSKLPEQ ERNYNLGMSETLKTLALGCHYGISDEHAE DKVKKMRLPRNYOLTSGYR-PAMDLSFIKLT PSGEAM-VDLABAR-BOYS GTGER-RUFFREKKLAENHIEL-WYM KIELGWQYGP-VRDDNKRQHFCL-VFFSKLPEQ ERNYNLGMSETLKT-TOR-TOR-TOR-TOR-TOR-TOR-TOR-TOR-TOR-TO							
EPAPARRROTHADGGEGEDEJGHTDDEVY LQCTATHIKEQOKLCLAAGGGRAICLESTS NSKNYPPDLSICTFYLEQSLSYRALQEMLANT VESSEQVDVEWKMPMATAGGGGRATILL YGHAILLRISYSGMYLCLSTSSSSTIKLAFD VOLQEDTTGEACWWINPASKQRSGEKYR VQDDLLLVSVSSERYLHLSYGNOSLHYDAAF QQTLWSVAPISSGSRAAQOYLIGGDYRLH GRMDECLTYPSGEBIGEGRATYHYEGGAVS VHARSURELTRAV WSGSHIRWOOFPRIR HYTTGKYLSLMEDKPILLJADKEKADYKSTIA FIFRSSEKLDVGVYREKEVDGKMGTSEKYGDS VCYIQHYDTGLWLTYQSVDVKSVRMGSIQR KAMPHIEGHMDDGISLSSGHEESKTARVIRS TYFLFRYRRIGLALSKKAASTVDLFESYSL SLQDLIGYTHPPDEHLEHEDKGNRLRALKNR QNFQEEGMMINLVELDIRLHYSSGAAFATD VAGREAGESWKSILNSLYELLAALIRGNKN GAGFSGSLDWLISRLELBASSGLEVLHGVL VESPEALMIKEGHKSIISLLDKHGRNIKVLD VLSLCVCHGVAVRSNQHLICDNLLPGDLL LQTRLVNHYSSMRPINFLGVSEGSAQYKKWY YEMNVDHTEPPVTAEATHLRVGWASTEGYSP YPGGGEEWGGNOVGDDLFSYGFDGLHLWSG CARTVSSPNQHLLETDDVISCCLDLSAPSISF RINGQPVQGMFENFINDCLFFFVVSFSAGIKV RFLIGGRIGEFKLIPPOYAPCYBAVLPKKL KVEHSREYKQBRYTTDLLGFTVSLTQAAFT PPVDTSGVIPPHERERREKLAENHIELWWM KIELGWQYGPVRDDNRQHFCLVEFSKLPEQ ERYNNLQMSLETLKILLALGCHGUSDEHAE DKVKKMKLPKNYQLTSGYRPAPMDLSFIKLT PSGEAMYDKLABAHNVARDRRGGWTY GIQQDVKNRRPRRVYPTPLDDRTKKSNKDS LRAVRTLLGGGTNYGBEHAE DKVKKMKLPKNYQLTSGYRPAPMDLSFIKLT PSGEAMYDKLABAHNVARDRRGGWTY GIQQDVKNRRPRRVYPTPLDBTKKSNKDS LRAVRTLLGGGGNTAGCGWTD KARWHGNBFRLYPYTPLDBTKSSNKDS LRAVRTLLGGGGNTAGCGWTD SEPHENIEVTRUGGTDSSCLAKFOPLOG KARWHGNBFHYGSSWQAGDVCCMVDM NEHTMMFTLNGELLDBGSGSAGTNAKGVT GGREVGROPPANTNRDTMWS.KRLPQDLOY PSHEHEIVTRUGGTDSSCLAKFOPLOG KARWHGNBFHYGSSWQAGDVVCCMVDM NEHTMMFTLNGELLDBGSGSAGTRORDDY PSHEHEIVTRUGGTDSSCLAKFOPLOG GFPVCSLGVAQVGRNNGGKDVSTLKYFTC GLQEGVSRPAGNOPDELGSGERAFROOLOG KARWHGNBFHYGSSWQAGDVVCCMVDM NEHTMMFTLNGELLDBGSGEAFRENOOLOG GFPVCSLGVAQVGRNNGGKDVSTLKYFTC GLQEGVSRPAGNOPDELGSGERAFRENOOL SIRGNSTYNYSVRIFGQERANNWGWTIS DFIQTYDTGFOLDBKNTYTVTVILGGERGKHYE SIRRSNCYMCAGESMSFQGGRNNNGLBIGC VVDAASGLITFANACKELSTYYQVETSFTLIPP AVFAQATSBNYFQFELGRIKNYMPLSAGLFKS EHNDYPQCPFLHAVPOLLSNLWSMRMOLEIG KYDNELSTTGELTGELLSKHYTYNLTIPT EFTKSTILTPDSNKKHGLPGIGLSTSLRPRMQF RVDNLSERGGWILVQCLDFLOGMSLHHEEN RVDNLSERGGWILVQCLDFLOGMSLHHEEN RSDNLSHETGERLSTYFLOFT FOR THE THE TOTAL THE TYPOTTS TORT THE THE TYPOTTS	552	1902	Α	4197	2	14302	ARPPPAPGSRQQKQKAAPGAAAAAELRGAR
NSKNYPPDLSICTTVLEOSISVRALQEMLANI VEKSEGQVDVEK WFMMKTAQGGGRITLI YGHALLRINYSGMYLCCLSTRSSTDKLAPI VGLQEDTTGEAC WTHPASKQRSGEKVR VGDDLILVSVSSERYLHLSYGMGSLHVDAAF QOTLWSVAPISSGEAAQGYLIGGDVLRLH GIMDECLTVPSGEHGEEORRTVHYEGGAVS VIARSLWELETLAVAWSGHRWGOPFRLR HYTTOK YLSLMEDKNILLIMDKEKADVKSTA FFFRSKEELDWGVKEVDOMGTSEKYGDS VCYIQHVDTGLWLTVQSWVESVRMGSIQR KAIMHEGHMDDGISLSRSQHEESRTARVIBS TVFLFNFRIRGLDALSKKAKASTVDLIESVSI. SLQDLGYFHPPDEHLEHEDKONRLRALKOR QNLFQEEGMINLVLECDRLHVYSKAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN. CAQFSGLDWLISLERLEASSGELEVLHCVL VESFEALNIKEGHKSIISLLDKHGRNIKKVLD VLCSLCVCHGVAVRSOHLDCINLPGDLL LQTILVSHVSSMRPNIFLCVSEGSAQYKKWY YELMWJOHTEPFVTAEATHLRVGWASTEGYSP YFGGGEWGGGVGDDLFSYGFDGLHLWSG CLARTVSSPNQHLRTDDVISCCLDLSAPSISF RINGQPVQGMFENNIDGLFPPVVSFSAGIKV KVEHSREYKQERTYTRDLLGPTVSI.TQAAFT PPVDTSQVI.PPHERREKLAENHELWVMN KIELGWGYGFVRDDMKRGPPCLVEFSKLFEQ ERNYNLQMSLETKTLLALGGTVGSLT SEGMYDKLERHELWMN KIELGWGYGFVRDDMKRGPPCLVEFSKLFEQ ERNYNLQMSLETKTLLALGGTVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFKLT SEGAMYDKLEARHNVWARDERRQGWTY GIQDVVRRNRNRLYPTTPLDDRTKKSNKDS LREAVRTLLGYGVILLARGROCK GGRFRIFTRAEKTYAVKAGRWYFEFETVA GGRFRIFTRAEKTYAVKAGRWYFEFETORD GFFRYDLEARGRAFTRAGDDAY GFFRYNDEGARAFTWORDHNINGLIGC VVDAASGLTFIARGKAGRUNGLINGLICC VVDAASGLTFIARGKAGRUNGLINGLICC VVDAASGLTFIARGKAGRUNGUNGLINGL							EPAPARRRGTMADGGEGEDEIQFLRTDDEVV
NSKNYPPDLSICTTVLEGSLSVRALQEMLANI VEKSEGQVDVEK WKFMMKTAQGGGHRTLI YGHALLRINSYGMVLCLSTSRSSTDKLAFF VGLQEDTTGEAC WVTHPASK, QREGEK VR VGDDLILVSVSSERYLHLSYGMGSLHVDAAF QOTLWSVAPINSGEAAQOYLIGGDVLRLH GHMDECLTVPSGEHGEBORRTVHYEGGAVS VHARSL WKLETLKAV AWSGHR WGOPPLR HYTTGK YLSLMEDKNILLIMDKEK AD VKSTA FFFRSKEKLDVGVKLKEVDGANGTSEK YGDS VCYNGNYDTGLWLTVQSVDVKSVFMGSIQR KAIMHEGHMDDGISLSRSQHEESRT AR VTRA FFFRSKEKLDVGVKLEVDGANGTSEK YGDS VCYNGNYDTGLWLTVQSVDVKSVFMGSIQR KAIMHEGHMDDGISLSRSQHEESRT AR VTRA VGREAGESWKSILNSLYGHLESRT AR VTRA VAGREAGESWKSILNSLYSLLAALIRGNRKN AGREAGESWKSILNSLYSLLAALIRGNRKN AGREAGESWKSILNSLYSLLAALIRGNRKN AGREAGESWKSILNSLYSLLAALIRGNRKN AGREAGESWKSILNSLYSLLAALIRGNRKN VGLYCHOWAYSSMRPNIFLCVSEGSAQYKKWY YSLMWYDHTEPFYTAATHLR VGWASTEGYSP YFGGGEBWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNOHLHRTDDVISCCLDLSAPSISF RINGQPVQGMPENNIDGLFPPVVSFSAGIKV VRIEGREGWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNOHLHRTDDVISCCLDLSAPSISF RINGQPVQGMPENNIDGLFPPVVSFSAGIKV VRIEGRGFGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG							LOCTATIHKEOOKLCLAAEGFGNRLCFLESTS
VEKSEGQVDVEKWKFMMKTAQGGGHRITL YGHALIRINSYGMYLCLISTSRSTDKLAFD VGLQEDTTGEACWWTHPASKQRSGEGEV VGDDILLVSVSSEFYLHLSYGMCSILHVDAAF QQTLWSVAPISGSEAAQVLIGGDVLRLH GIMDECLTVPSGEHGEORRTVPGOAVS VHARSLWRLETLRVAWSGSHIRWGQPFRLR HYTTGKYLSLMEDKYLLIMDKEKADVKSTA FIFRSSKEKLDVGWFREVDMGTSEIKYGDS VCYQHVDTIGLWILTVQSVDVKSVBMGSIGR KAMHHEGHNDDGISLSRSQHEESKTARVIRS TVFLFNRFIRGLDALSKKAKASTVDLFIESSYL SLQDLIGYHPPDEHLEHEDKQNRIKALKNR QNLFQEEGMINLVLECIDELHVYSSAAHFAD VAGREAGESWKSINSLYELLAARINR CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKGENHESSISLLKMERNKNL CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKGENHENISISLLKMERNKNLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LITRLVNHVSNRAPINLGVSEGSAQVKKWY YELMYDHTEFFVTABATHLRVGWSTEGYSP YRGGGEWGGNGVGDDLFSYGFOLHWSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSIS FINGQPYQGMEPHTIGGLFPVSLTGAAFT PEVDTSQIVLPHLERIBEKLAENHELWVMN KIELGWQYGFVRDDNKRQHPCLVEFSKLPC ERNYNLQMSLEFILKTLLALGCRYGSDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRKLIT PREDEAMYDKLAERAHNVWARRENGOWTY GIQQDVKNRNRNRLVPTYTLDDRTKKSNLDS LREAVRTLLGYGNLEFROEDHARACVCS GTGERFREFRAEKTYAVKAGRWYFEFETVTA GDMRWGWSRGCOPDQLGLSSERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVCGMVDW NEHTMAFTLNGGILDDSSGELAFKDPVGD GFPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLGRGYFPADNTRDITMDLTKKSNLDS NEHHBYTRIDGTIDSSECLATVAGGRGAG LFGFKNDLEDYDDSDFEVLMTTAHGHLV PROFINGERFFFLYDYCGROWD STIDIMFYRLSMPGLAFRAFDDVGD GFPYCSLGVAQVGRMNFGKDVSTLKYFTIC GLGRGYFPAPNTRDITMDLTKKSNLDS NINDIMFYRLSMPGLAFRAFDDVGD FMGTSTYTYSVRIPEGGEFANNVGUNGTS DFHOYDTGFDLDRVRTNTTVTLGDERGKVM PRINEHHBYTRIDGTIDSSECLATVAGGRGAG LFGFKNDLEDYDDSDFEVLMTTAHGHLV PROFINGENEHGVRSWGAGDVVCGMVMD NEHTMAFTLNGGILDDSSEGLAFKDPDVGD GFPYCSLGVAQVGRMNFGKDVSTLKYFTIC GLGRGYFPAPNTRDITMDLTKKSNLDS NINDIMFYRLSMPGCAEVSKTVYAGGLPGGAG LFGFKNDLEDYDDSDFEVLMTTAHGHLV PROFINGENEHGVTRIDGTIDSSECLKYTGSGFGGAG LFGFKNDLEDYDDSDFEVLMTTAHGHLV PROFINGENEHGVTRIDGTHDSSECLKYTGSGFGGAG NINDIMFYRLSMPGCAEVSKTVYAGGLPGGAG LFGFKNDLEDYDDSDFEVLMTTAHGHLV SKRSNCYYAVGAGSMSFGCGRNNNQLEIC VVDAASGLITFANGKELSTTYQVGPSTTLP AVPAQATSTNYGFELGRIKNVMFLSACLTKS ENCHOPPOQCPGRIFGGGASMSFCGGGRNNNGLEGE SSFSYSISNECOYGSFEPPOLLKSKTJOALTE	ſ	(					
YOHALLRISYSGMYLCLSTRSSTDKLAP VGLODTIGEACWTHPASKQRSGEKYR VGDDLILVSYSSERYJHLSYGMGSLHVDAAF QOTLWSVAPISSGEAAQOYLIGGDVLRLIH GIMDECLTVPSGEHGEEORRTVHYEGGAVS VHARSLWRLETLRVA WSGSHIRWGPFRLR HVTTGKYLSLMEDKNLLIMDKEKADVKSTA FTFRSSKERLDVGWRKEVDOMGTSERYGDS VCYIQHVDTIGNLTYQSVDVKSYRMGSIGR KAIMHEGHMDGGISLSKGHEESKTARVIRS TVFLFNEFRGLALSKKAKASTVDLFIESVSL SLQDLIGVFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLGUDRLHVYSAAHFAD VAGREAGESWSILNSLYBLLAALIRGNRNN CAQFSGSLDWLISRLERLEASSGELEVLHCVL VESFEALNIKEGHKSIISLLDKHGRNIKVLD VLCSLCVCHGVAVRSNQHLICDHLFORDLL LQTRLVNHVSSMRPNIFLGVSGSAQYKKWY YELMMJDHTEPFLYAATHLRVGSGAQYKKWY YELMMJDHTEPFLYAATHLRVGSGAQYKKWY YELMMJCHTSSPROHAATHLRVGSGAQYKKWY REMGORVQQMEENFINDGLFFVGFDGLHLWSG CLARTVSSPNQHLRTDDVISCCLDLSAFSISF RINGQPVQGMEENFINDGLFFPVSLTQAAFT PIPVDTSQIVLPHLERIEBKLAENHELWNM KIELGGWYGGFVEDDNKRCHPCLFFSKLFPQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMLKPKYQLTSVYRPAMDLSFIKLT PSQEAMYDKLAENAHNVARDERQGWTY GIQDDVKNRRWPLLYTPLDDRTKKSNKDS LREAYRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWFGNDSTLKYT GIQDVKNRRWPRLYTPTLDDRTKKSNKDS LREAYRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWFGNDSTLKYTFIC GDMRVGWSRPGCCPDQGLGSBERAFAEDG KAQRWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG SAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPDGGLGSBERAFAEDG GAGWYGWSRPGCCPGGLGSBERAFAEDG GAGWYGWSRPGCCPGGLGSBERAFAEDG GAGWYGWSRPGCCPGGLGSBERAFAEDG GAGWAGAGGAG GAGWAGATSPRWYGGGAGAAG GAGWAGAGGAGAGAGAGAGAGAGAGAGAGAG	1	ŀ					
VGLQEDTTGEACWWTHPASKQRSGEGEKY VGDDLLVSVSSERYLH.SYGNGSLHVDAAF QQTLWSVAPISGSEAAQGYLIGHYGGAVS VGARSLWRLETLRVAWSGHTRWQOPFIR. HYTTGKYLSLMEGHGEDGRRTYGHYGGAVS VHARSLWRLETLRVAWSGHTRWQOPFIR. HYTTGKYLSLMEGHGEDGRRTYGHYGGAVS VHARSLWRLETLRVAWSGHTRWQOPFIR. HYTTGKYLSLMEGHGEDGRRTYGHYGGAVS VGARSLWRLETLRVAWSGHTRWQOPFIR. HYTTGKYLSLMEGHGEDGRRTYGHYGGAVS VGYGYMEGHDELORDKOMGTSERY QDD VCYIQHYDTGLUHYTOSVDVSYBMGSIQR KAIMHEGHMDDGISLSRSQHEESRTARVTRS TYPLFYRRTRGLDALSKKAKASTYDLPIESSYS SLQDLIGYHHYPGEHLEHDEDKONRIKALKSVIR. QNLFQEEGMINLVLECIDRIHYYSSAAHFAD VAGREAGESWKSINSLYELLAALIGNRKN CAQFSGSLDWLISKLERLEASSGLEVLHCVL VESFEALNIKGGHKSISILLCHKRIKKVLD VLCSLCVCHGVAVRSNQHLICONLLPGRDLL LQTRLVYHYSSAMPHICLOVSEGSAQYKKWY YELMYDHTEPFYTABATHLRVGWASTEGYSP YNGGGEWGGRGYGODLFSYGFOLHLWSG CLARTVSSNQHLLRTDDVISCCLDLSAPSISF RINGOPYQGMEDFYNTIOLLFPTVSLTQAAFT PRODTSQIVLPHLERIERKLAENHELWVAM KIELGGRYGGFVRDFINGDLFFPVSLTQAAFT PPVDTSQIVLPHLERIERKLAENHELWVAM KIELGWYGFVRDDVKRCHPCLYSTALDAFT PPVDTSQIVLPHLERIERKLAENHELWVAM KIELGWYGFVRDDVKRCHPCLYFSLAPG ERNYNLOMSLETLKTLALACCHVGISDEHAE DEVKKKMKLPKNYQLTSYKYRAPDLSFIKLT PSQEAMYDKLAENAHVWARDRIRGGWTY GIQQDVKNRRNRRLYFTTLDDRTKKSNKOS LREAVRILLOYSVLAENAHOLSFIKLT GDMRYWGSRGCOCPOQLGISSERAFAFDOF KAQRWHQGNEHYGRSWQAGDVVCCMVDM NEHTMAFTLINGGILDDSSGELARKDPOVGD GFTPVCSLGVAQVGRMNFGKDVSTLKYTTIC GLGRGYFPFANTNEDTIMVLKRRLPGFLQV PSNHEHIEVTRIDGTIDSSELARKDPOVGD GFTPVCSLGVAQVGRMNFGKDVSTLKYTTIC GLGRGYFPFANTNEDTIMVLKRRLPGFLQV PSNHEHIEVTRIDGTIDSSELARKDPOVGD GFTPVCSLGVAQVGRMNFGKDVSTLKYTTIC GLGRGYFPFANTNEDTIMVLKRRLPGFLQV PSNHEHIEVTRIDGTIDSSELAKTDPUXGD RGCRYPFFANTNEDTIMVLKRRLPGFLQV PSNHEHIEVTRIDGTIDSSELAKTDPUXGLEGGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLUP DFHQYDTGFDLDRVKTYTYTLGDEGGKVHE SIRRSNCYMYCAGEMSSPGGGRNAVWGANFISCH PHQYDTFGPLDRVKTYTYTLGDEGGKVHE SIRRSNCYMYCAGEMSSPGGGRNAVWGANFISCH SIRRSNCYMYCAGEMSSPGGGRNAVWGANFISCH RVDULETTEGELLKFHYHTHLRYSMPOLIE CKYDVSRISERGGWLVQCLDPLQFMSLHIFEEN RYDLILLTEGEGLKFHYHTHLRYSMPOLIE EKKSITLFPDENKKHGLPGGLSTSLRPRMQFL KYDVSRISERGGWLVQCLDPLQFMSLHIFEEN SSPSYSISNECOYSTSCOYSEPPOLLLKSKTIQALTE							
VGDDLL/SVSSERY-ILLSYONGEH-WDAFF QOTL-WSVAPISGESGRAQGY-LIGIDVLRILH GIMDECLTYPSGERIGEORRTVHYEGGAVS VHARSI-WRLETLR/A-WGSSHIR-WGOFFRLR HYTTGKYLSLMEDKY-LLMDKEKADVKSTA FTFRSSKEKLDWGYRER-WGOMGTSEIKY-GDS VCYQH-VDTGL-WLTYQSVDWKSVRMGSIQR KAIM-HEGHMDDGISLSRSQ-HEESTLARVIRS TVFLFNRFIRGLDALSKKAK-STVDL-PIESVSL. SLQDLIGYTHPPDEHLEHEDKQNRLRALINR QNLFQEEGMINU-LECIDRL-HYSSA-AHFAD VAGREAGESWKSILNSL-YELLALIRGNRN CAGFSGSLDWIJSRLEPLEASSGLIE-VHCVL VESPEALNIKEGHIKSIISLLDKHGRNIKN CAGFSGSLDWIJSRLEPLEASSGLIE-VHCVL VESPEALNIKEGHIKSIISLLDKHGRNIKN-UD VLCSL-CV-HGY-WAYSN-GHLIGH-VHSC-WL- VESPEALNIKEGHIKSIISLLDKHGRNIK-VL- UCSL-V-HGY-WAYSN-GHLIGH-VHSC-WL- VESPEALNIKEGHIKSIISLLDKHGRNIK-VL- UCSL-CV-HGY-WAYSN-GHLIGH-VHSC-WL- VESPEALNIKEGHIKSIISLLDKHGRNIK-VL- VESPEALNIKEGHIKSIISLLDKHGRNIK-VL- UCSL-CV-HGY-WAYSN-GHLIGH-VS-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- YELM-WHTE-PUT-AE-ATHL-WA- KELGGRGEFKL-PP-POYAR-CY-A-VI-YEKL KVEHSREY-KGER-YTHR-LIG-LIW-SG CLART-YSS-PW-HD-LSFRIKT- PIP-VDTSQ/VL-PH-LERIREKL-AENHEL-WWNN KELGGWGGP-WORDDIN-RQ-HC-U-FFSKL-PEQ EEN-YH-LOG-WHTE-WA- KKKLL-RW-NQ-LTSG-YK-P-PM-LSFRIK- PIP-VDTSQ/VL-PH-LERIREKL-AENHEL-WWNN KELGWGG-WYDDDIN-RQ-HC-U-FFSKL-PEQ EEN-YH-LOG-WGWS-RG-CQ-PD-QEL-GS-BEA-PA-FD-GF KAQR-WH-GG-RK-YN-V-C-W-JUM- NEHT-MM-FTL-NGEIL-LD-SGS-BL-AFKD-FD-GF KAQR-WH-GG-RK-YN-V-W-G-W-G-W-JUM- NEHT-MM-FTL-NGEIL-LD-SGS-BL-AFKD-FD-GF KAQR-WH-GG-RK-YN-V-W-G-W-G-W-M-M- NEHT-MM-FTL-NGEIL-LD-SGS-BL-AFKD-FD-GF KAQR-WH-GG-RK-YN-V-W-G-W-G-W-M- NEHT-MM-FTL-NGEIL-LD-SGS-BL-AFKD-FD-GF KAQR-WH-GG-RK-YN-V-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-	1	1					
QQTLWSVAPISGSERAQGYLIGGDVLRLLH GHMDECLTYPSGGHSEGERFUPYFOGAYS VHARSL WRLETLRVAWGSHIRWGOPFILR HVTTGKYLSLMEDKYLLLMDKEKADVKSTA FFFRSSKEKLDVGVRGEVDGMGTSEIKYGDS VCYQHVDTGLWLTYGSVDKSVRMGSIQR KAIMHEGHMDGISLSRSQHEESRTARVIRS TYFLPNRFRGLDALSKKAKASTVDLPESYSL SLQDLIGYTHPPDERLEHEBKOKQNRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAHFAD VAGREAGESWSKSINSLYSLALALIRGNKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBALARGNKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHKISSISLLBARGNKNA VERMYSSENGHLENTOLSCAPALSSGILEVLHCVL VESSELVHYSSENGHLENTOLSCAPALSSGILEVLHCVL VESSELVHYSSENGHLENTOLSCAPALSSGILEVLHCVL KVEHSREYKQERTYTRDLLGPTVSLTOAATT PROTTERPTVTSALATHARVAN KELGGWGPGPFVESRCHACHANHELWAMN KELGWGPGPVRDDNKRQHPCLVEFSKLPQ ERNYNLQMSLETKLTLLALGCHSDSHELWTHARVAN KELGWGPGPVRDDNKRQHPCLVEFSKLPQ ERNYNLQMSLETKLTLLALGCHSDSHELWTHARVAN KELGWGPGPPAVDLSRKHLT PSQEAMVDLKARTHARVANARDRINGGSBERAFAFDGF KAQRWHQGNEHTAMTANDARAFETSTATA GDMRVGWSRPGQPDQELGSDERAFAFDGF KAQRWHQGNEHTYGNVQCMTDM NEHTMAFTLNGEILLDBGSSELAFKDPDWD GFIPVCSLGVAQWGNAFTGEKTVAYKAGRHANGELIG GGGYEFALTRAKKTPYSTSHSARLTEDVLADDRDDY DFLMGTSTTYYSVRIPFOGEPANVWGWITS DFHQTDTGFDLDRVRTVTVTLCDEKGKWHE SIKRSNCYMVCAGESMSPOGPANVWGWITS DFHQTDTGFDLDRVRTVTVTLCDEKGKWHE SIKRSNCYMVCAGESMSPOGPLLWSTLIPPEN NFDQFDLRWFTLANGELSTYQVEFSTLLEP AVFAQATSPHVYFOFELRIKHNLWFLYPMT BETKSTILFPDENKHALDGIGSTSLRRRMOF SYSPISISMECYGYSPEFLULBKSTTIQMLE	1						
GHMDECLTYPSGEHGEEQRITYHYEGGAYS VHARSL WRLETLR VAWSGIRIR WGOPFRLR HYTTGKYL SILMEDKNILLMMKER ADVKSTA FFFRSSKEKL DVGVYRKE VDGMTSELKY GDS VCYIQHYDTGL WLTYQS VDVKSVRMGSIQR KAMMHEGHMDDGISLSRQHEESTRARVIRS TVFLDNRIFIGLDALSKKAKASTYDLPESVSL. SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINU VLECIDRILYSKAHFAD VAGREAGESWKSILNSL YELLAALIRGNRKN CAQFGSGLDWI, SILENLERSCHEV, HCVL. VESPEALMIKEGHIK SIISLLDKHGRNIK VLD VLCSLCVCHGVAVRSNOHLICDNILPGRDLL LQTRLVNRVSSMRPNIFLOVSEGSAGYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSE YRGGGEEWGGNOVODLIESVGFOGLHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAFISIF RINGQPVQGMFEVNINDGLFPVVSFSAGIKV RFLLGGRIGGFKFLPPPGVAFCYFAVLPKEL KVEHSREVKQERTYTRDLLGPTVLTQAAFT PIPVDTSQIVLPPHLBRIREKLAENHHELWVAN KIELGWQYGPVRDDNKRQHFCLVEFSKLPEQ ERNYNLQMSLETIKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKFAPMDLSPIKLT PSQEAMVDKLAENAHNWARDRIRQGWTY GIQQDVKNRRNPRLVPYTHLDDRTKKSNKDS LREAVRTLLGGYNLEAPDARARACVCS GTGEFFRRAEKTYAVRAG WYFEFETVTA GGMRVGWSREGCOPDGLE GERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMAMFTLNGEILLDDSGSLAFKDFDVOD GFIPVCSLGVAQORRNNFGKDVSTLKYFTIC GLQGYGFFAVNTNRDITMWLSKRLPQTQVQ PSNEHEHEVTRIDGTIDSSSCLAFKDFDVOD GFIPVCSLGVAQORRNNFGKDVSTLKYFTIC GLQGYGFFAVNTNRDITMWLSKRLPQTQVQ PSNEHEHEVTRIDGTIDSSSCLAFKDFDVOD GFIPVCSLGVAQORRNNFGKDVSTLKYFTIC GLQGYGFFAVNTNRDITMWLSKRLPQTQVQ PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV STRIPEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQVQV PSNEHEHEVTRIDGTIDSSSCLAFKDFTQLAFK PAYAGATSPMVYGFELGRIKNMFLSTYLPT AVFAGATSPMVAHALCSHUZSHAHRWPESTYPMT BETKSTILTFPDENKKHOLPGIGSTSLRRRMQFI KVDVARASGLATFIRDFLARKYTAWMESTYPMT BETK	1						
VHARSL WRLETLR VAWSGSHIR WGGPFRLR HVTTGK YLS IMEDENNILL MEDENNILL MEREADVKSTA FFFRSSEKLD VGVREV DGMGTSEIK; GDS VCYIQHVDTGLWLTY QSVVKSVRMGSIQR RAMMHEGHMDDGISLSRSQHEESR TARVIRS TVFLFNRFRGLDALSKKAKASTVDLPESVSL SLQDLIGYHFPPDEHLEHERGKQNRLRALKNR QNLFQEEGMINLVLECIDRLHYYSSAHFAD VAGREAGESWSKIINSLYELLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHIKSTISLLAKIRGNIKNL UVCSLCVCHGWAVRSNOHLICIDNLLPGDLL LOTRLVNHVSSAMFRNFIG (VSGSAQYK WY YELMVDHTEPFVTAE ATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFOGLH WSG CLARTVSSNOQHLRTDDVISCCLDLSAPSISF RNOQPVQGMESINNIDGLFFPVVSFSAGIKV RFLLGGRRIGEFRLPPPVAFAVCYAVLPKEKL KVEHSREYKQERTYTRDLIGFTVSLTQAAFT PPVDTSQNLPPHLERIREALARNHELWWN KELGWYGPVRDDNKRQHFCLVFFSKLFQ ERNYLOMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVKLKARLBKNYQLTSGYKPAPMDLSFIKLT PSQEAMVKLKARLBKNYQLTSGYKPAPMDLSFIKLT PSQEAMVKLKARLBKNYQLTSGYKPAPMDLSFIKLT PSQEAMVKLKARLBKNYQLTSGYKPAPMDLSFIKLT GGMRVGWSRPGQPDQELGSDERAFAFDGF KAQRWHQGNEKHTYANKAGNYFFFTVTA GGMRVWSRRGCQPDQELGSDERAFAFDGF KAQRWHGGNEHYGSWQAGNYGCMYDM NEHTMMFTLNGEILLDDSGSELAFKDFDVG GFFYCSLGVAQVGRMNFGKULVTLKYFTIC GLGGYGFFANTNRDITMWLSKLPQDLQV PSNIEHEEVTRIGTIDSSPCLLAVTGKSFGSON SYTDIMFYRLSWFECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDREATTRPFTNNRIDYAGEKFSILKP AVFAQATSPMYYGFERNIKUPA SIKRSNCYMVCAGESMSPQGRNNYGLEG SIKRSNCYMVCAGESMSPQGRNNYGLEIG KYKLMSTARPFICARDSTRANGFISHCH SIKRSNCYMVCAGESMSPQGRNNYGLEIG SIKRSNCYMVCAGESMSPQGRNNYGLEIG VVDAASGLITFIANGKELSTYYQVEFSTLIPP AVFAQATSPMYYGFELKIRNAMPLSHIPEEN RSVDILELTEQELLKFHYHTLRLYSAVCALG NIRNAHALCSHVOPGILVAENNYMPLSAGLFKS EKKNPYPQCPFRIHVOPGISHVLWSRMPOELL KVDVASRISERQGWIN VGCLIPP LOPANSLHIPEEN RSVDILELTEQELLKFHYHTLRLYSAVCALG NIRNAHALCSHVOPGISHVLWSRMMPOELL KVDVLSRISERGGWIN VGCLIPP LAKKHRICHT ETKSTILTFDENKKHGLEGIGTSTSLRFRMQFI SSPSYSISNECYGYSPEFULLKSKTIQMLTE	1						
HVTTGKYLSLMEDKYLLLANDKEKADVKSTA FFTRSSKEKLDVGVRKEVGMGTSEIKYGDS VCYQHVDTGLWLTYQSVDVKSVRMGSIQR KAIMHEGHMDDGISLRSQHEAVERSTARVIRS TVFLFYRFIRGLDALSKKAKASTVDLPIESVSL SLQDLIGYFHPPDEHLEHBCKQNRLRALKNR QNLFQEEGMINLVLECIBLYVSSAAHFAD VAGREAGESWKSLINSLYELLAALIRGNKN CAQFSGSLDWLISLIELELSSGLEVLHCVL VESPBALNIIKEGHIKSIISLLDKHGRNHKVLD VLCSLCVCHOVAYRSNHCLDOLLPORDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTABATHLRVGWASTEGYSF YPGGGEWGGNGVGDLDLFPVVSFSAGIKV RILLGGRHGEFKFLPPGVSPCYGAVLYREKL KVEHSBEYSGENTYTRDLIGFFPVVSFSAGIKV RFLLGGRHGEFKFLPPGVAPCYEAVLYREKL KVEHSBEYSGERTYTRDLIAGGHVGISDEHAE DEVKKMKLPKNYQLTSGYRAPMDLSFIKLT PROGAWDKLAENAHNVANDAKHFLAGGHVGISDEHAE DEVKKMKLPKNYQLTSGYRAPMDLSFIKLT PSQEAMVDKLAENAHNVARGRWYFEFTVTA GQQDVKNRRNPRLVPYTPLDDEHKKSNKDS LREAVETLLGGYNLLBARDRRQCWTY GQQDVKNRRNPRLVPYTPLDDEHKKSNKDS LREAVETLLGGYNLLBARDRRQCWTY GGGERRFRAEKTYAVKAGRWYFEFTVTA GDMRVGWSRPGCQPDQDGDHAARAEVCS GTGERRIFRAEKTYAVKAGRWYFEFTVTA GDMRVGWSRPGCQPDQDGDHAARAEVCS GTGERRIFRAEKTYAVKAGRWYFEFTVTA GDMRVGWSRPGCQPDQBGBERAPAFDGF KAQRWHQGNEHYGRSWQAGDVGCMVDM NEHTIMMFILNGEILLDBGSELAFKDDVGD GFFPVCSLGVAQVGRMNNGENSTLKYFTIC GLQEGYEPFANTINRDITMWLSKALDFIQUY PSNHEHEVTRIDGTIDSSGERAFADFOR KAQRWHQCNEHYGRSWQAGDVGCMVDM NEHTIMFTLNGEILLDBGSELAFKLOFLQV PSNHEHEVTRIDGTIDSSGENAFARDFOR KAQRWHQCNEHYGRSWQAGDVGCMVDM NEHTIMFTLNGEILLDBGSELAFKLOFLQV PSNHEHEVTRIDGTIDSSGERAFANDFUNDDY DFLMQTSTYYYSVRFPQGEPANVVVGWITS DFHQYTSJSMPRCAEVFSKTVAGGLPGAG LIGFRNDEAGLYNGTRJGKLSTYVQVEPSIKLPP AVFAQATSSNVVGGESSNRSQGRNNINGLEIGC VVDAASGLLTFANGKELSTYYQVEPSIKLPP AVFAQATSSNVVGGESSNRSPGGRNNINGLEIGC VVDAASGLLTFANGKELSTYYQVEPSIKLPP AVFAQATSSNVVGGESSNRSPROGL KVDVSRISERGGWLVQCLDFLOFMSLHIPEEN RSVDILLITEQEELLKFHYHTLRLYSAVCALD NHRVAHALCSHVDEPQLLYAIENKYMFRIGHES RSVDILLITEQEELLKFHYHTLRLYSAVCALD NHRVAHALCSHVDEPQLLYAIENKYMFRIGHE RSVDILLITEQEELLKFHYHTLRLYSAVCALD NHRVAHALCSHVDEPGLLKSKTIQMLTE							-
FIFRSSEKLDVGVKEVDGMGTSEIKYODS VCYGIQHVDTICLWITYQSVKSVRMGSIQR KAIMHHEGHMDDGISLSRSQHEESRTARVIRS TVFLFYNRFIRGLDALSKKAKASTVDJEESVSL. SLQDLIGYTHPPOHLEHEDKQMRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAHFAD VAGREAGESWKSLINSLYSLISLALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPRALMIRGCHIKSISLYBLIAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPRALMIRGCHIKSISLYBLIAGNIRGNKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPRALMIRGCHIKSISLYBLIAGNIRGNKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPRALMIRGCHIKSISLYBLIAGNIRGNKN CAQFSGSLDWLISRLEGVSESGGILEVHCVL VESPRALMIRGCHIKSISLYBLIAGNIRGNKN VLOSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRFNIRGVSEGSAGIKV YELMVDHTEIPFVTABATHLRVGWASTEGYSP YPGGGEEWGGNGVODDLFSVGFDGLILWSG CIARTVSSPNOHLIRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFVVSSAGIKV RFLLGGRHGFEKFLPPPOTAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PPVDTISQIVLPFHLBRREKLAENHELWWMN KELLGWYGPVQDDNKRGPCFVFTSALFQE ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPNYQLTSOPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRRQGWTY GIQODVKNRENPRI LYPTIDLTKKSNDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERRIFRAEKTYAKAGRWYFEFTVTA GDMRVGWSRRGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGMVDM NEHTIMFTLINGEILLDDSSELAFKDFDVGD GFIPVCSLGVAQVGRMNGKDVSTLXYFTIC GLOGVEPPAVNTNEDITIMULSKRLPOPLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVTSKTVAGGLPGAG LGGYGPPAVNTNEDITIMULSKRLPOPLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVTSKTVAGGLPGAG LGGYGNDLDVDSDSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVTSKTVAGGLPGAG LGGYGNDLDVDSTSDFPLMKTAHGHLVP DRVDKDKEATKYPTYYTYTUTTUGDEKGKVHE SIKRSNCYMVCAGESMSPQGGRNNNGLEIGC VVDAASGLLTFANGKELSTYYQVEFSKLPP AVPAQATSSPNVOFELGRKNNMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRNIPQEL KVDNSSISERGGWLVQCLIGKNNMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRNIPQEL KVDNSSISERGGWLVQCLIGKNNMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRNIPQEL KVDNSSISERGGWLVQCLIGKSKTIPQHE SSPSTYSINSCKCYOYSEPTHLAENKYMPGLLC NIHVAHALCSHVDPQLLKKSKTIQMLIE	1						· · · · · · · · · · · · · · · · · · ·
VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR KAIMHHEGHMDDGISLSRQHEESRTARVIRS TVPLFNRFRGLDALSKKAKASTVDLPIESVSL SLQDLIGYTHPPDEHLEHEDKQRRKALKNR QNLFQEEGMINLVLECIBRALKNR QNLFQEEGMINLVLECIBRALKNR VAGREAGESWKSILNSLYELLAALIRGNKN CAQFSGSLDWLISLLERLESGLEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNIKVLD VLCSLCVCHGVAVRSNQHLICDNLJCRDLL LQTRLVNHVSSMRPINLGVSEGSAQYKKWY YELMVDHTEPFVTAFATHLRVGWASTEGYSF YPGGGEWGGNGVODDLFGYSGFGGLHLWSG CLARTVSSPNOHLIRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLFKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PPVDTSQIVTPPHLERIREKLAENIHELWYMN KELGWQYGPVRDDDNKRQHPCLVEFSKLPFQ ERNYNLQMSLETIKTLLAIGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWANDRIRGWTY GIQQDVKNRNPRLVPYTLDDRTKKSNNDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFTPVCSLGVAQVGRMNFGKDVSTLYFTIC GLQGYEFPAVNTNRDITNWLSKKLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSON SNTDMFYRLSMPECAEVTYQVSFRSILYFT DFVQNDLEDYDADSDFEVLMKTAHGLPV DDRVDKNKARTHENPLADRDDY DFLMGTSTYYYSVRIPFGQEPANVWWGWITS DFHQYDTGFDLDRVRTLYTICDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNIGLEIGC VVDAASGLLTFANGELLSTRILDVLADDRDDY DFLMGTSTYYYYSVRIPFGQEPANVWWGWITS DFHQYDTGFDLDRVRTLYTICDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNIGLEIGC VVDAASGLLTFANGKELSTYVQVEPSTKLFP AVFAQATSSNVPQFELGRIKNNMPLSAGLFKS EHKNPVPQCPPRLHVQLSHVLIKSKTIQMLIFE RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALGSHVDEPGLIKSKTIQMLIFE RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALGSHVDEPGLIKKSKTIQMLIFE RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALGSHVDEPGLIKKSKTIQMLIFE RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALGSHVDEPGLIKKSKTIQMLIFE RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALGSHVPDEPGLIKKSKTIQMLIFE RSSPOLSINGERCYVYSEPTILKSKTIQMLIFE RSSPOLSINGERCYVYSEPTILKSKTIQMLIFE	-						-
KAMMHEGHMDDGISLSROHEESRTARVISE TVPLFNFERGLOALSKAKAS TVOLPESYSL SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLECIDRLHYVSSAAHFAD VAGREAGESWKSLINSLYSLYSLLAALIRGNRKN CAQFSGSLDWLISBLERLEASSGILEVHCVL VESPEALMIKEGHKSSILDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPINLE GSSAQVKK WY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDLFSYGFGDLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENTNIDGLFFYVYSFAGIKV RFILGGRHGEFKFLPPPOTAPCYFAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTOAAFT PIPVDTSQIVLPPHLERREKLAENIHELWVMN KIELGWYGPVRDDNKRGPCLVEFSKLFQ ERNYNLQMSLETILKTLLALGCHVGISDEHAE DKVKKMKLPSNYQLTSGYAPMDLSFKLT PSQEAMVDKLAENAHNVWARDRINGGWTY GIQDVKNRRNPRLVPYTLDDRTKKSINDS LREAVRTLLGYGYNLEAPDQDHAARAEVGS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDGLGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTIMMFTLINGEILLDSSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLYFTIC GLQGYEPFAVNTNEIDTMWLSKRLPOFLQV PSNHEHIEVTRIDGTIDSSPCLKVTOKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGLPGAG LFGFKNDLEDVDADSDEPLJMKTAHGHLVP DDRVDKDKEATKFEFNNHKDYAQEKFSRLKQ RFILARTKFDYSTSHASALTEDVLADDRDDY DFLMGTSTYYYSVRIFFGQEPANVWGWITS DFHQYDTGFBLDRVBTLYTLOBEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFANGKELSTYYQUEFSTKLIP AVPAQATSBNYFQFELGRIKNVMPLSACLFKS EHKINVPQCPPRLHVGLSKYLWGMLIFEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKIFSTEGELSTYYQUEFSTKLIP AVPAQATSBNYFQFELGRIKNVMPLSACLFKS EHKINVPQCPPRLHVQFLSKYLWGRIFGEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKNKAMPISACLFKS EHKINVPQCPPRLHVQFLSKYLWGRILFEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKKNKAMPISACLFKS EHKINVPQCPPRLHVQFLSKYLWGRILFEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKNAMPLSACLFKS EHKINVPQCPPRLHVQFLSKYLWGRILFEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKANGPGLIR RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKANGFILENKAMPGLIR RSVDILELTEQEELLKKHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKANGFILENKAMPGLIR RSVDILELTEQEELLKKHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKANGFILENKAMPGLIR RSVDILELTEQEELLKKHYHTLRLYSAVCALG NIRWABALGSKYDFQCJRSKIKANGFILENKYMPGCLIR RGTSSTTILFTDDTNKKKHGLFGIGLSTSLRP	1						
TVFLFNRFIRGLDALSKKAKÁSTVDLPIESVSL SLQDLIGYHPPDELLEIDKONRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGLDWLISRLERLEASSGLEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD WLCSLCVCHGVAVRSNGLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFDGLH.WSG CIARTYSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVGGMFEVNTDGLFFPVVSFSAGIKV RFLLGGRHGEFKLPPPGVAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PPVDTSQTVLPHLERIEKLAENHIELWVMN KIELGWQYGFVRDNKRQHPCLVFFSKLPEQ ERNYNLQMSLETIKLTLALGCHVGISDEHAE DKVYKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHWARDRROGWTY GIQQDVKNRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERTRIFRAEKTYAVKAGRWYFEFETTVTA GDMRVGWSRPGCOPDQELGSBEAFATDGF KAQRWHQGNEHYGRSWQAGDVGGMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFFVCSLGVAQVGRNMFKGKDYSTLKYFTIC GLOEGYSFFAVNTNDTIMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVYQKSFGSQN SNTDIMFYSLSMPECAEVPSKTVAGGLPGG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAGERPSRLKQ RFLJRRTKPDYSTSHSARLTEDVLADDRDDY DFLMGTSTYYYSVRIFPQGEPANVWVGWITS DFHQYDTGFFDLDRWTTVLIGDEKGKYHE SIKRSNCYMYCAGESMSPGGRNNINGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLPP AVPAQATSPNVPQFPELDRLKSKTUNGPLLER EIKSNCYMYCAGESMSPGGRNNINGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLIPP AVPAQATSPNVPQFPELDRLMSKRTUNFLSAGLFKS EHKNPVPQCPPRLHVQFLSHNLWSRMPNQFL KVDVSRISERGGWLVQCLDPLQFMSLHPEEN RSVDILELTEGEELLKFNYHTLRLYSAVCALG NHRVHAHALCSHVDEPJCLJKRSNYMPGLR AGYYDLLDHLSSYATARLMMNNEYVPMT EETKSTILTPDENKKHGLPGIGLSTSLRPRMQF SSPSSYSISISNECYQYSPEPLDLIKSKTUNGLIEG	1						
SLQDLIGYFHPPDEHLEHEDKONRIRALKNR QNLFQEEGMINLVLECIDLHVYSSAAHFAD VAGREAGESWKSII.NSLYELLAALIRGYRKN CAQFSGSLDWLISRLERLEASSGIEVLHCVL VESPEALINIKEGHIKSILDKHGRIHKIKVLD VLCSLCVCHGVAVRSNQHLICDNLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CIARTVSSPNQHLLISTDDVISCCLDLSAPSISF RINGQPVQGMPENFNIDGIFPFVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLGPTVSLTQAAFT PIPVDTSQIVLPHLERIREKLAEHIELWVMN KIELGWQYGPVDDNKRQHPCLVEFSKLFQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRRQGWTY GIQQDVKNRRNPRLVPYTFLDDDTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFREFAAEKTYAVKAGRWYTEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGGMVDM NEHITMMFTLNGEILLDDSGSELAFKDFDVGG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKYTQKSFGSQN SNTDIMFYRLSNPIECAEVPSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATTYPETNIKDYAQEKPSILKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTDTGFDLDRVRTIVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPQGGRNNGGLEIGC VVDAASGLITTIANGELSTYYQVEPSTKLPP AVFAQATSPRVFQFELGRIKNVMPLSAGLFKS EHKNPVPCCPPRLHVQFLSHTVUNRSRMPNGEL KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALIG NHRVAHALCSHVDEPQLJKRSKHMPGELIR AGYYDLLIDHLSSYATARLMMNNEYVPMT EETKSTILTPDENKKHOLPGIGLSTSLRPRMQF							
QNIFQEEGMINLYLCIDRIHYVSSAAHFAD VAGREAGESWSLINSLYELLAALIRGNIKNLO CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHIKSIISLLOKHGRNHKYLD WLCSLCVCHOVAVRSNICHLOONLICONLIPGRDLL LQTRLVNHVSSMRYNIFLOVSEGSAQYKKWY YELMYDHTEPPVTAEATHLRUGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CIARTYSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFILLGGRHGEFKLPPPOYAPCYFAVLPKEKL KVEHSREYKOERTYTROLGFTVSLTOAAFT PIPVDTSQIVLPHLERIREKLAENHELWYMN KIELGWQYGPVRDNKRQHPCLVEFSKLPQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRKLT PSGEAMYDKLAENAHWAARDRRGGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYTEFETVTA GDMRVGWSRPGCQPDQELGSDERAFATDGF KAQRWHQGNEHYGRAYGAGDVVGGMYDM NEHTMMFTLNGGILLDDSGSELAFKDFDVG GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEFFAVNTNEDITIMVLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYSLSMPELGAEVFSKTVAGGLPAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRTVDKDKEATKPEFTVTLGDEGKEAFFSKTVAGCLPAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRTVDKDKEATKPEFTVTLGDEGKEYFSKLVQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMGTSTYYYSVRIFPQGEFANVWVGWITS DFHQYDTGFFDLDRVTTVLGDEKGKVHE SIKRSNCYMVCAGGSMSPGGGRNNNGLEIGC VVDAASGLITFIANGESTYYQVEPSTKLPP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISSERGGWLVQCLDPLQFMSLHPEEN RSVDILELTEGEELLKFHYHTLRLYSAVCALG NHRVHAHALCSHVDEPGLJKKSKTIQHTER							
VÄGREÄGESWKSILNSLYELLAALIRGNIKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNIHKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPILGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFYVSFSAGIKV RFLIGGRIGEFKFLPPGYAPCYEAVLPKEKL KVEHSREYKGRTYTRDLLGPTVSLTQAAFT PIPPVDTSQIVLPPHLERIREKLAENIHELWVMN KIELGWQVGPVADDNERQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKKMLPKNYDLSGYKAPMDLSFRLT PSQEAMVDKLAENAHNWARDRIRQGWTY GIQQDVKNRRIPRLVPYTTLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIRRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCOPELGSDERAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSLAFKDFDVGD GFDVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEFFANTHRDITMUSKRLPGFLQV PSNEEHEVTRIDGTISSPCLKVTQSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGFKNDLEDVDADSDFEVLMKTAHGHLVP DRVDKDKAEATKFFFNNHKDYAGEFPSILKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRLVTVTUTLGDEKGKVHE SIKRSNCYMVCAGESMSPGGRNNNGLEIGC VVDAASGLITTIANGKELSTYYQVEPSTKLFP AVFAQATSPNYFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGLLYTHERLYRMPGLK KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGVLYAIENKYMPGLK KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGLLYTAMERVMPGLK KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGVLYAIENKYMPGLK KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGVLYAIENKYMPGLK KVDVSRISERQGWVVQCLDPLQFMSLHPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPGVLYAIENKYMPGLK RGYDDLJDHILSSYATARLMMNNEYIVPMT EETKSTILFPDENKKHGLPGIGLSTSLRPRMQF SSSFSVSISNECYQVSEFFIDLLKSSTIQMLTE	[						
CAQFSGSLDWLISRLERLASSGILEVLHCVL VESPEALNIKEGHIKSISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICDNLPGRDLL LQTRLVNHVSSMPNIFLGVSGSGAQYKKWY YELMVDHTPFFYTAFATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CIARTVSSRNQHLLRTDDVISCLDLSAFSISF RINGQPVQGMEENFINDGLFFPVVSFSAGKV RFLLGGRIGGEKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTEDLLGPTVSLTQAAFT PIPVDITSQIVLPPHLERIREKLAENIHELWVMN KIELGWQYGPVRDDNRRQHFCLVEFSKLPEQ ERNYNLQMSLETIK.LALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNUWARDRRQGWTY GIQDDVKNRRIPRLVPYTFLDDRTKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFAEKTYAVXAGRWYFEFETVTA GDMRVGWSRFGCQPDQELGSDERAFATDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPGFLQV PSNEEHEVTRIDGTIDSSTCLKVTQSFGSQN SNTDIMFYRLSMPECAEVPSKTVAGGLPGAG LFGFKNDLEDVDADSFEVLMKTAGFLLV DRYNKDKEATKPFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMGTSTYYYSVFRGQEPANVWGWITS DFHQYTGFDLDRVRTVTTLGDEKGKVHE SIKRSNCYMCAGESMSPGQRNNNGLEIGC VVDAASGLLTFLANGKELSTYYQVEPSTKLFP AVFAQATSPNNFQFELGRIKNVMFLSAGLFKS EHKNPVPQCCPPRLIFYLSHVLSWENDFOLL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTILLYSAVCALG MIRVAHALCSHVDGYLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTILLYSAVCALG MIRVAHALCSHVDGYLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTILLYSAVCALG MIRVAHALCSHVDGYLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTILLYSAVCALG MIRVAHALCSHVDGLISTSLRRMMPS	•		ľ				QNLFQEEGMINLVLECIDRLHVYSSAAHFAD
VESPEALNIKEGHIKSIISLLÜKHGRNHEVLD VLCSLCVCHGVAVRSNQHLICDNILLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGCDDLFSYGFGELH.WSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIFLGLFFPVVSFSAGIKV RFLLGGRIGEFKFLPPGYAPCYBAVLPKEKL KVEHSREVKQERTYTRDLLGPTVSLTQAAFT PPVDTSQIVLPPHLEREKLAENHEHLE WYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKKMKLPKNYGLSGYKPAPMDLSFELT PSQEAMVDKLAENAHNWARDRIRQGWTY GIQQDVKNRRNPRLYPTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRFGCOPELGSDERAFATDGF KAQRWIGGNEHYGRSWQAGDVVCGMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFDVCSLGVAGVGRMNFGKDVSTLKYFTIC GLQEGYEFFAVNTNRDITMWLSKRLPGFLQV PSNEEHHEVTRINGEINSPCLKVTQSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDVJADSDFEVLMKTAHGHILVP DRVDKKDKEATKFFFNNHKDYAGEFPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMGTSTYYYSVETPGQEPANWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS CHAPAUCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFILNFGQEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLITINGLGLESTYYQVEPSTKLFP AVFQQATSPNVFQEELSTYYQVEPSTKLFP AVFQQATSPNVFQEELSTYYQVEPSTKLFP AVFQQATSPNVFQEELSTYYQVEPSTKLFP AVFQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHPEEN RSVDILELTEGEELLKFHYHTILLYSAVCALG MHVAHALCSHVDEGLISTSLRFRMQF SSSFSVSISNECYQVSPEFFLDLLSKKTIQMLTE	1 .						VAGREAGESWKSILNSLYELLAALIRGNRKN
VLCSLCVCHOVAYRSNOHLICDNILPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNOHLRTDDVISCCL DLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGVAPCYEAVLPKEKL KVEHSREYKQGERTYTRDLLGFTVSLTQAAFT PPVDTSQJVLPPHLERREKLAENIHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFKLT PSQEAMYDKLAENAINVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTFLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCOPDQELGSDERAFAPDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECASFYSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEPNHKDYAQEKPSRLKQ RFLLRRTKFDYSTSHSARLTEDVLADDRDDY DFLMGTSTTYTYSVRIFFGQEPANNWVGWITS DFHQVDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPQGRNNNGLEIGC VVDAASGLLTFLANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPYPQCPPRLHVOFLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLR.YSAVCALG NIRVYAHALCSHLKPHYHTLR.YSAVCALG NIRVYAHALCSHLKFHYMFGLLR AGYYDLLDIHLSSYATARLMMNNEYIVPMT EETKSTTLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFYSISNECYQYSPEPPLDLIKKSKTIQMLIE						•	CAQFSGSLDWLISRLERLEASSGILEVLHCVL
LQTRLVNIHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNQHLLRTDVISCCLDLSAPSISF RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PPVDTSQVLPPHLERREKLAENHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPNGVLTSGYKPAPMDLSFEKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRNPRLVPYTLDDRTKKSNKDS LREAVRTLLGGYGYNLEAPDQDHAARAEVCS GTGEFFIFFAEKTYAVKAGRWYFFFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEFFAVNTNRDITMWLSKRLPQFLQV PSNIEHHEVTRIDGTIDSSFCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAGEKPSRKQ RFLLRTKPPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQVDTGFDLDRVRTVTVTLGDERGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSSRNVFQFELGRIKNVMPLSAGLFKS EHKNPYPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLR;VSAVCALG NHRVAHALCSHVDEPQLLYAENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSTTLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNEC(YQYSPEPPLDLKSKTIQMLIE							VESPEALNIIKEGHIKSIISLLDKHGRNHKVLD
YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNQHLLRTDDVISCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQGERTYTRDLLGFTVSLTQAAFT PPVDTSQIVLPPHLERIREKLAENIHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFBLLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDDMFYRLSWPIECAEVFSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQVDTGFDLDRVRTVTVLLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSTILFPDENKKHGLPGGLSTSLRFRMQF SSPSFYSISNECYQYSPEPPLDLKSKTIQMLIE							VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL
YPGGEEWGGNGVGDDLFSYGFDGLH.WSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RNNGQPVQGMENFNIDGLFPFVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERITYTRDLLGPIVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN KELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRPPLVPYTPLDDRTKKSNKDS LREAVRILLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFFTVTA GDMRYGWSRGCOPDQELGSDERAFADGF KAQRWIQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPECAEVPSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKAEATKEPFNHKDYAQEKPSRLKQ RFLLRRTKFDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPQGGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPMOFL KVDVSRISERQGWLVQCLDPLOFMSLHIPEEN RSVDILLELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLY ALENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFFLUDILKSKTIQMLTE				[			LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY
CIARTYSSPNQHLIRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHEL WVMN KEELGWQYGPVRDDNKRQHPCL VEFSKLFEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMYDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLBAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYTEFETIVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSSGELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFVLMKTAHGHLPP DRVDKDKEATKPEFNNHKDYAQGKPSPIKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFFGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTJLGBEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNGLEIGC VVDAASGLITHANGKELSTYYQVEPSTKLFP AVFAQATSPNVFGELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMFNQFL KVDVSRISERQGWLVQCLDPLOFMSLHIPEEN RSVDILETTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLL AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFILDILKSKTIQMLTE	i l						YELMVDHTEPFVTAEATHLRVGWASTEGYSP
CIARTYSSPNQHLIRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHEL WVMN KEELGWQYGPVRDDNKRQHPCL VEFSKLFEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMYDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLBAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYTEFETIVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSSGELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFVLMKTAHGHLPP DRVDKDKEATKPEFNNHKDYAQGKPSPIKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFFGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTJLGBEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNGLEIGC VVDAASGLITHANGKELSTYYQVEPSTKLFP AVFAQATSPNVFGELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMFNQFL KVDVSRISERQGWLVQCLDPLOFMSLHIPEEN RSVDILETTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLL AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFILDILKSKTIQMLTE							
RINGQPVQGMENFNIDGLFFPVVSFSAGIKV RFLLGGRHIGEFKILPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN KELGWQYGPVRDDNKRQHFCLVEFSKLJFEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAINVWARDRRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRYGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFTPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPAVNTNRDITMWLSKRLPOPLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFSVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYQVEPSTKLFP AVFAQATSPNVFOFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF	1						
RFLLGGRIGEFKFLPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PPPVDTSGJVLPPHLERIREKLAENHIELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLMSLETLKTLLALGCHYGISDEHAE DKVKKMKPNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRYGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLMGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYSVRIPFQGEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQRNNNGLEIGC VVDAASGLLTFIANGKELSTYQVEFSTKLFP AVPAQATSPNVFOFELGRINVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLDHILSSYATARLMMNNEYIVPNT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF	1						
KVEHSREYKQERTYTROLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIPFGGEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLIR AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSTILFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECVQYSPEFPLDILKSKTIQMLTE							
PIPVDTSQIVLPPHLERIREKLAENIHEL WVMN KIELGWQYGPVDDNKRQHPCLVEFSKLPEQ ERNYNLOMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEPTSNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERGGWLVQCLDPLQFMSLHIPEEN RSVDLELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSTITJPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	]						l
KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETILKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTGSYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPOFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVPSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFFGQEFANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVAGAESMSPGGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDLELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE						·	
ERNYNLÓMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCOPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNIEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGKNDLEDVDADSDFFVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLIRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLGYMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARIMMNNEYTVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	ł i			ł	ľ		•
DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFADTOGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNEHHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLIRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPCCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEGEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLIR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF							
PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPFAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYYSVRIFFGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKKSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEFQLLYAENKYMPGLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1						
GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCOPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFTPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFFGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMOF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE				1			•
LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGEFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMOF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE					j	l	
GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSFNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE							
GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLWQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	]		- 1	ļ			
KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFLANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVOILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			[	ĺ		ļ	
NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE				İ			
GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			- 1				
GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	[		- (	Í		i	
PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	]			J	J		
SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	[			l	1	ł	
LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	]		i	l			
DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE				1	1		
RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	[	ĺ	- 1	- 1	ĺ	1	
RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		Į	İ	l	I	1	DRVDKDKEATKPEFNNHKDYAQEKPSRLKO
DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		1	1	l	I		
DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		l	·ł	ł	l	1	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			l		l	ł	, ,
VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			l	ļ		i	
AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1 1	ļ	ļ	1		ļ	
EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	]		l	l	1		
KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	}	1	ł	1			
RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		1	[			l	
NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			l			ì	
AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			-		ļ	1	
EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		-	. l	•	- 1	ļ	
SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			.			ļ	
				-	ł	ļ	
AVREGSLHARDPVGGTTEFLFVPLIKLFYTLLT		[	ſ	ı	1		
					1		AVKEUSLHAKDPVUGTTEFLFVPLIKLFYTLLI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isolaucine, K=Lysine, L=Laucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ĺ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ł			peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
	Ì					LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
						QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
1	•					VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
		•				ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
1						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
1						NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
					İ	DSKKSSTLQQLISETMVRWAQESVIEDPELVR
1	1					AMFVLLHRQYDGIGGLVRALPKTYTINGVSV
						EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
	·					LGDIMNNKVFYQHPNLMRALGMHETVMEV
1						MVNVLGGGESKEITFPKMVANCCRFLCYFCR
					·	ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
						STPLDVAAASVMDNNELALALREPDLEKVVR
1						YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
						YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
1						CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
1						GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
						SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
1						LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
1						AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
						EVGFLPDLRAAASLDTAALSATDMALALNRY LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
				"		VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
						MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
1				'		YERCWKYYCLPGGWGNFGAASEEELHLSRK
						LFWGIFDALSQKKYEQELFKLALPCLSAVAG
						ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ
						PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
						LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE
					ſ	KEIYRWPIKESLKTMLARTMRTERTREGDSM
1 1						ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
						NVTLSRDLHAMAEMMAENYHNIWAKKKKM
						ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
]			]		j	QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
						YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
						PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
	· .			j		SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
					ŀ	RISLFGNDATSIVNCLHILGQTLDARTVMKTG
]				-		LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNOPKGVTOIINYTTVALLPMLSSLFEHI
			1			GQHQFGEDLILEDVQVSCYRILTSLYALGTSK
			1			SIYVERQRSALGECLAAFAGAFPVAFLETHLD
						KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
						SLEKLMEEIVELAESGIRYTOMPHVMEVILPM
						LCSYMSRWWEHGPENNPERAEMCCTALNSE
, J		J	J	j		HMNTLLGNILKIIYNNLGIDEGAWMKRLAVF
		- 1	1	I	ĺ	SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
			ĺ	Ī	ĺ	SEEDHLKAEARGDMSEAELLILDEFTTLARDL
				1	ŀ	YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
1 1		į		1		MVAEVFIYWSKSHNFKREEONFVVONEINN
1	J	J	J	. ]	J	MSFLITDTKSKMSKAAVSDOERKKMKRKGD
1				ĺ	l	RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
			l	]	l	LAKNRFSLKDTEDEVRDIRSNIHLQGKLEDP
		1	İ	İ		AIRWQMALYKDLPNRTDDTSDPEKTVERVL
		l	ł	. 1	1	DIANVLFHLEQKSKRVGRRHYCLVEHPQRSK
	Į	ļ	1	1	j	KAVWHKLLSKQRKRAVVACFRMAPLYNLPR
	. [			ŀ	ł	HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
1			1	-2-	l	KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
					<u> </u>	EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						EEEVKSFEEKEMEKQKLLYQQARLHDRGAA EMVLQTISASKGETGPMVAATLKLGIAILNGG NSTVQQKMLDYLKEKKDVGFFQSLAGLMQS CSVLDLNAFERQNKAEGLGMVTEEGSGEKV
	٠.					LQDDEFTCDLFRFLQLLCEGHNSDFQNYLRT QTGNNTTVNIIISTVDYLLRVQESISDFYWYY SGKDVIDEQGQRNFSKAIQVAKQVFNILTEYI
						QGPCTGNQQSLAHSRLWDAVVGFLHVFAHM QMKLSQDSSQIELLKELMDLQKDMVVMLLS MLEGNVVNGTIGKQMVDMLVESSNNVEMIL KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK
	•					RDFHKAMESHKHYTQSETEFLLSCAETDENE TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH MPNDTRLQTFLELAESVLNYFQPFLGRIEIMG SAKRIERVYFEISESSRTQWEKPQVKESKRQFI
						FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA QISESDLNERSANKEESEKERPEEQGPRMAFF SILTVRSALFALRYNILTLMRMLSLKSLKKQM
			7.			KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV ASVFRGFFRIICSLLLGGSLVEGAKKIKVAELL ANMPDPTQDEVRGDGEEGERKPLEAALPSED LTDLKELTEESDLLSDIFGLDLKREGGQYKLIP
						HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK EEEKEEKEETKSEPEKAEGEDGEKEEKAKED KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY QQKLLNYFARNFYNMRMLALFVAFAINFILL
						FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS SHRIIAVHYVLEESSGYMEPTVRILPILHTVISF FCIIGYYCLKVPLVIFKREKEVARKLEFDGLYI TEQPSEDDIKGQWDRLVINTQSFPNNYWDKF
	,		-1100	,		VKRKVMDKYGEFYGRDRISELLGMDKAALD FSDAREKKKPKKDSSLSAVLNSIDVKYQMW KLGVVFTDNSFLYLAWYMT
553	1903	A	4199	31	767	LPELNGRGAGLRRAEPSERGGGAERTQQVAA LPLSHGHSHGGGGCRCAAER/VGAARGSAAC AYGLYLRIDKGRLQCLNESREGSGRGVFKPW ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK
					,	GIIIMGEDDDSHPSEMRLYKNIPQMSFDDTER EPDQTFSLNRDLTGELEYATKISRFSNVYHLSI HISKNFGADTTKVFYIGLRGEWTELRRHEVTI CNYEASANPADHRVHQVTPQTHFIS
554	1904	A	4200	1	961	GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL EICIKACKNLAYGEEKKKKCNPYVKTYLLPD RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT
	130					WDFEDSTTQSFRWHPLRAKADKYEDSVPQS NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT
						LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV TPAQLRQSSLELTVWDQALFGMNDRLLGGT\ RLGSKGDTAVGGDACSQSKLQWQKVLSSPN LWTDMTLVLH
555	1905	A	4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH NHHENERRWQQERLHREEAYYQFINELNDE DYRLMRDHNLLGTPGEITSEELQQRLDGVKE
	,			-		QLASQPDLRDGTNYRDSEVPRESSHEDSLLE WLNTFRRTGNATRSGQNGNQTWRAVSRTNP NNGEFRFSLEIHVNHENRGFEIHGEDYTDIPLS DSNRDHTANRQQRST\SPVARRTRSQTSVNFN
L						GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNGI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end mucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  GGAAGIPRANASRTNFSSHTNQSGGSELRQRE GQRFGAAHVWENGARSNVTVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV QQTTRRSVRRRGRTRVFLEQDRERERRGTAY TPFSNSRLVSRTTVEEGEESSRSSTAVRRHPTTT LDLQVR\RIPPGENRDRDSIANRTRSRVGLAE NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN NRSQHREGSSQDRQAQGDSTEMHGENETTOP
556	1906	A	4212	3	462	HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH FFLLNESDDDDRIRGLTKEQIDNLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR
						KSPENTEGKDGSKVTKQEPTRRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRFSCLTLQTSWGHRHGPPRP\ANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKSSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFSVKMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP
559	1909	A	4235	ì	323	KFSIPFFLRWSFTLV\PRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS
560	1910	A	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPPDTWDSFPLFQ LLNDFLRTGLLICGNGK\FHKHLQDLFAPLVV R/YMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLHWPEEEFGK HLEQRLKLMASDMIESCVKRTR\IAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL\CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHTYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

SEQ ID NO: of NO: of NO: of nucl-   NO: of nucl-   cotide   seq-   uence   u	GLI LIFG EFQ DW QF PGK PSGPR PEV CLIF KLA ILPV WF ETG FAV PQR ILQ G
nucleotide bed bed in uncleotide docation orresponding to last amino acid residue of peptide sequence   914	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
cotide sequence  USSN 09/496 914  USSN 09/496 914  Beginning to first amino acid residue of peptide sequence  USSN 09/496 914  Beginning to first amino acid residue of peptide sequence  Beginning to first amino acid r	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
sequence    Sequence   1914   1946   1914   1946	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
uence  914	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
amino acid residue of peptide sequence    T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, Y=possible nucleotide deletion, \=possible nucleotide insertion   INTLEPVEEDAEMRGDQGTPTTGLLMIVLO; FMKGNTIKETEAWDFLLAL/GVYPTKKHLIF DPKKLITEDFVRQRYLEYRRIPHTDPVDVEF WGPRTNLETSKMKVLKFVAKVHNQDPKDV PAQYCEALADEENRARPQPSGPAPSS   S62	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
residue of peptide sequence    Y=Tyrosine, X=Unkmown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion   Possible nucleotide insertion	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
peptide   sequence	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
sequence    NTLEPVEEDAEMRGDQGTPTTGLLMIVLG    FMKGNTIKETEAWDFLLAL\GVYPTKKHLII   DPKKLITEDFVRQRYLEYRRIPHTDPVDYEF   WGPRTNLETSKMKVLKFVAKVHNQDPKDV    PAQYCEALADEENRARPQPSGPAPSS    S62	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
INTLEPVEEDAEMRGDQGTPTTGLLMIVLG) FMKGNTIKETEAWDFLLALIGVYPTKKHLII DPKKLITEDFVRQRYLEYRRIPHTDPVDYEF WGPRTNLETSKMKVLKFVAKVHNQDPKDV PAQYCEALADEENRARPQPSGPAPSS  562 1912 A 4260 I 1498 MVTWLYRFLPTSNMAAKLRSLLPPDLRLQI WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTIRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFTIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LRERQRMRNQLELAARGPLRQTFTHNPLI PGKDNPPNIPSSISSSSSKPKSKYPWHDTLG FGS 1913 A 4265 623 116 MGGLAPTQTLEFTNREYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
FMKGNTIKETEAWDFLLALIGVYPTKKHLIE DPKKLITEDFVRQRYLEYRRIPHTDPVDYEF WGPRTNLETSKMKVLKFVAKVHNQDPKDV PAQYCEALADEENRARPQPSGPAPSS  562 1912 A 4260 1 1498 MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNILLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPTVREYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	QF GF GPSQ GPSQ GLI CLIF KLA ILPV WF ETGV FOR ILLQ G
DPKKLITEDFVRQRYLEYRRIPHTDPVDYEF WGPRTNLETSKMKVLKFVAKVHNQDPKDV PAQYCEALADEENRARPQPSGPAPSS  562 1912 A 4260 I 1498 MVTWLYFFLPTSNMAAKLRSLLPPDLRLQI WHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIEAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEFTVREYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	QF QF QK PSQ SGPR PEV CLIF KLA ILPV WF ETG FAV PQR ILLQ G
WGPRTNLETSKMKVLKFVAKVHNQDPKDV PAQYCEALADEENRARPQPSGPAPSS  562 1912 A 4260 1 1498 MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITIMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	QF QF PSQ SGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
PAQYCEALADEENRARPQPSGPAPSS  562 1912 A 4260 1 1498 MVTWLYRFLPTSNMAAKLRSLLPPDLRLQI WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASGG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSSSSSSKPKSK YPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	QF PGK PSQ GGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
562 1912 A 4260 I 1498 MYTWLYRFLPTSNMAAKLRSLLPPDLRLQE WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS'SSSSSSKPKSK YPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	GK FPSQ SGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
WLHARLQKCFLSRGCGSYCAGAKASPLPGI MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	GK FPSQ SGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
MAMGLMCGRRELLRLLQSGRRVHSVAGPS WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSISSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	SPSQ SGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
WLGKPLTTRLLFPAAPCCCRPHYLFLAASG SLSTSAISFAEVQVQAPPVVAATPSPTAVPE ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSISSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	GGPR PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
SLSTSAISFAEVQVQAPPVVAATPSPTAVPE' ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAETV VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	PEV GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
ASGETADVVQTAAEQSFAELGLGSYTPVGL QNLLEFMHVDLGLPWWGALAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAETV VQSSDLQWMRNVIRMMPITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	GLI CLIF KLA ILPV WF ETG FAV PQR IRQ LLQ G
QNLLEFMHVDLGLPWWGAIAACTVFARCL PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAETV VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	CLIF KLA ILPV WF ETG FAV PQR TRQ LLQ G
PLIVTGQREAARIHNHLPEIQKFSSRIREAKL GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	KLA ILPV WF ETG FAV PQR FRQ LLQ G
GDHIEYYKASSEMALYQKKHGIKLYKPLILI TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSSISSSSSKPKSK YPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	ILPV WF ETG FAV PQR FRQ LLQ G
TQAPIFISFFIALREMANLPVPSLQTGGLWW QDLTVSDPIYILPLAVTATMWAVLELGAET VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	WF ETG FAV PQR FRQ LLQ G
QDLTVSDPIYILPLAVTATMWAVLELGAETVQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	ETG FAV PQR FRQ LLQ G
VQSSDLQWMRNVIRMMPLITLPITMHFPTA FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	TAV PQR TRQ LLQ G
FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQ VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	PQR TRQ LLQ G
VVHDLDKLPPREGFLESFKKGWKNAEMTR LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	TRQ LLQ G
LREREQRMRNQLELAARGPLRQTFTHNPLL PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG  563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	LLQ G
PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSSKPKSKYPWHDTLG   FGKDNPPNIPSS\SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	G `
563 1913 A 4265 623 116 MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQ THGTRRTLSSGPSNNLPLPLSSSATMPSMQQ	
THGTRRTLSSGPSNNLPLPLSSSATMPSMQC	EON
1	QCK
	<b>V</b> L
PRGSGNPPHGTSILTAPPALLPHPPTHPTQQS	
LIQENNNTNHTHSHTHTYTETLSFFLYICVN	
DRMEWGKSVF	
564 1914 A 4270 3 368 ILKRKLSSLNSEVSTIONTRMLAFKATAOLF	LFIL
GCTWCLGLLQVGPAAQVMAYLFTIINSLQC	
FIFLVYCLLS\QQVQKQYQKWFREIVKSKSE	
ETYTLSSKMGPDSKPSEGDVFPRTSE	
565 1915 A 4288 83 406 RNSRPLWCSPPASQPRQAPVSQSCCCPLPSS	2222
PPSALLAPTKPRALGTLRLYECSPELGTTML	
PAWLLMLCQAPRPQDPDRLTQPEKSLQEA	
GOTGASRTPRT	L/u
	ΖЛ
GVRYLTLTFTCSTPWAESSTKFRHHMYTNV	
GLTSFGEKVVEELNRLGMMIDLSYASDTLIF	
VLEVSQAPVIFSHSAARAVCDNLLNVPDDII	
LLKKNGGIVMVTLSMGVLQCNLLANVSTV	
DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLV	
DVSTYPVLIEELLSRSWSEEELQGVLRGNLL	
VFRQVEKVREESRAQSPVEAEFPYGQLSTSC	SCH
FHLGASEWTPRLLIWR	
567 1917 A 4299 1 1106 GATPLGSVGGRTGKMDAATLTYDTLRFAEI	
DFPETSEPVWILGRKYSIFTEKDEILSDVASR	SRL
WFTYRKNFPAIGGTGPTSDTGWGCMLRCGG	:GQ
	YFS
VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ	
WYGPNTVAQVLKKLAVFDTWSSLAVHIAM	
NTVVMEEIRRLCRTSVPCAGATAFPADSDRI	
CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLT	
INEAYVETLKHCFMMPQSLGVIGGKPNSAH	
FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDE	
FHCQHPPCRMSIAELDPSIAVVRGGHLSTQA	
GAECCLGMTRKTFGFLRFFFSMLG	Λιπ.
<del> </del>	776
	, v 3
LMSVLIPKLPQLHGVRIFGINKY	mv.
569 1919 A 4302 186 531 WTFCLFL/WWVPESARWLLTQGHVKEAHR	

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT VTESKLEAEGKTKEKAREKERKKKS
570	1920	A	4308	3		RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS GKRNKLRVYYLSWLRNKILHNDPEVEKKQG WTTVGDMEGCGHYRVVKYERIKFLVIALKSS VEVYAWAPKPYHKFMAFKSFADLPHRPLLV DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE KEHREFRAKTNRDLEIKDQEIEKLRIELDESK QHLEQEQQKAALAREECLRLTELLGESEHQL HLTRQEKDSIQQSFSKEAKAQALQAQQREQE LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT KLKEECCTLAKKLEQISQ
572	1922	A	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE DFPETSEPVWILGRKYSIFTEKDEILSDVASRL WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ MIFAQALVCRHLGRDWRWTQRKRQPDSYFS VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD NTVVMEEIRRLCRTSVPCAGATAFPADSDRH CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLYT DINEAYVETLKHCFHGWPQFFG/VVHREGK PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
573	1923	A .	4333	363	1066	GGVPVGLASKPFQILYGHTNEVLSVGISTELD MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS LFLTIPNLAISWEGHIVVYSSTEEKTTLKVERM HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD FSKGSK
574	1924	A .	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK LPSARAKIRITSSPIFITFYILVFVVALVGIARA VVSMTVSTSNAATVADKILWEITRFFLLAIEL SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV TQGTLEILYPDAHLSAEDFNIYGHGGRQFWL VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY VYAGILALLNLLQGLGSVLLCFDIEGLCCVD ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	Ā	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP QDQDKLSKEDVLSFIQMHRA
576	1926	A	4365	69	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS EPLPLGGIRPTPGLEPKGRDLM

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	durce	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uence		1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		٠.	1	residue of		
					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	1	Ι.	peptide	i	/=possible nucleotide deletion, \=possible
	1000		1000	Sequence		nucleotide insertion
577	1927	A	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
	1					ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
		<u> </u>				FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	Α	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
						SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
						SGWSRTPDLR
579	1929	Α	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
1		1	1	[	1	FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
			l			CWPGWSSTPDLK
580	1930	Α	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
	i	i			İ	\VFKKGNHILHELFQNKEEGAFPNS/FYEASFT
			1			LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
						QLKSSDL .
581	1931	Α	4414	670	3	VLVHROCGGILRLRRKEAVSVLDSADIEVTDS
					i -	RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
		1				RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
1		ł				RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD
1		1				LVLHINGESTOGLTHAOAVERIRAGGPOLHL
İ		1				VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
		1				DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
1	•	ŀ				SPE
582	1932	A	4424	194	449	VLYIRKKKRLEKLRHOLMPMYNFDPTEEODE
362	1932	^	4424	174	447	LEQELLEHGRDAASVQAATSVQAMQGKTTL
1	{	1		'		PS\QGPLQRPSRLVFT\DVANAIHV
583	1933 .	_	4435	1	166	
363	1935 .	Α	4433	1	100	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
1						PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
584	1934		4439		(20	SAPPALLQDTSV
284	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
			1			APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
1						APATQHSQAGPATGQAYGPHTYTEPAKPKK
1						GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
						ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
						GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
	1005		116			QGPHGKAAQGGAAGAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
						SIFDDFAHYEKRQ
586	1936	A	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
						FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
						INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
					.	TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
				]		FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
1	•					LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
1						PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
						ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
				ĺ		LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
				l		PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
					· ;	LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
				ļ		SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
	•					FRAPPAINARLPFNFFFPFV
587	1937	A	4471	614	387	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
""	.,,,	4.	4471	7.7	207	CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
						NKHEPPCPATKRRHSGSI
588	1938	A	4480	1720	1450	
1 900	1230	A.	4460	1/20	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
1					1	CPANFCIII/DFLVETGFHHVGQASHELLTSGD
500	1020		4405	000	200	PPTSASQSAGITGMSYHTWFGES
589	1939	Α	4487	922	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
						PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
					İ	REAAQAMYIDSYNSRGFMINGNRVLGPCALL
Ll			L			PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide seq-		in USSN	nucleotide location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
1				amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	sequence	= 12 yrosine, X=Onknown, 1=Stop codon, 1=possible nucleotide deletion, 1=possible
				sequence		nucleotide insertion
						VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
						DTPNACATFNFLCHEGRVTGAALIPPPGGTSL TSLGQAAQ
590	1940	A	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
						PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
						RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
						VPHQGGLPGPIRVAPSSAGQREASQGPPGR
591	1941	A	4495	1444	1116	IAARFTLAKTWNQLKRP\TMIDSIKKTR\YIYT
						MEYYADTERNEIMSF\AGTWVELEAIILSKLM
			l			LKDNWVEDTIPQGAVPCTATAEGMKRLLFAL EPWDSSCFPHPSSGV
592	1942	Α	4496	2	919	RTRPLFSGRPTRPVCTMSDERRLPGSAVGWL
		[		_		VCGGLSLLANAWGILSVGAKQKKWKPLEFL
						LCTLAATHMLNVAVPIATYSVVQLRRQRPDF
						EWNEGLCKVFVSTFYTLTLATCFSVTSLSYHR MWMVCWPVNYRLSNAKKQAGHTVMGIWM
						GSFILSALPAVGWHDTSERFYTHGCRFIVAEI
						GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
						AVQVGRQADHRAFTVPTIVVEDAQGKRRSSI DGSEPAKTSLQTTGLVTTIVFIYDCLMGFPVL
						GPFSLADTHLSDLPYTWGDRDSGGACVM
593	1943	Α	4506	2	193	FFFEAESCSVPQAGVQRPDLGWLHAPPP\GSC
						HFPASASQVAGTTHARHHTQLIF\AFLVENGL C
594	1944	Α	4507	1327	647	KMAGGVRPLRGLRALCRVLLFLSQFCILSGG
						ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS
						CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ
				-		RYPANCTVR\DHVHCLGNRTFPKMLYCNWT
						GGYKWVYGLWLLRHHPRWGLGADRF\YLGP
						VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG
595	1945	A	4512	533	264	PADGSLYI FFFKMESYSVARLECSGAISAPCNLHLLGSNN
	1545	^	45.12	333	204	SPASASRV/AGNIGARHHTQQIFVLLVQMRVH
						YVGQDGLDLL/NLMIHPPRSPKVLGLQA
596	1946	A	4513	3	1674	HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV
						NLMLELLVQKKKQLEAESHAAQLQILMEFLK
	·					VARRNKREQLEQIQKELSVLEEDIKRVEEMS
						GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP
						PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS
						KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV
1						FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
						VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGORSKVYQEHEKRCWSV
					-	DFNLMDPKLLASGSDDAKVKLWSTNLDNSV
						ASIEAKANVCCVKFSPSSRYHLAFGCADHCV
		4				HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN
						EKNFV\GLASNGDYIACGSENNSLYLYYKGLS
				, , ,		KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV
100	1046		4516	50.6	804	CWRALPDGESNVLIAANS\QGTI\KVLELV
597	1947	Α.	4518	536	824	RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE
						QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP
598	1948	Α	4524	1	384	FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF
			L			TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID	SEO ID	Met	SEQ .	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	HOU	in NO:	nucleotide		F=Phenylalanine, G=Glycine, H=Histidine,
		1			location	
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
į	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			ì	peptide	•	/=possible nucleotide deletion, \=possible
ļ	ļ	ł	J	sequence	ļ	nucleotide insertion
	<del> </del>	<del> </del>	<del> </del>	Soquenee		RRREMQSQSVMLALRRGDAVWLLSHDHDG
	l	i	l			
	1		[	ł .		YGAYSNHGKYTTFSGFLVYPDLAPAAPPGLG
	1010	<u> </u>				ASELL
599	1949	Α	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP
[			1			NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG
l			İ		<u> </u>	VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK
į	l	ļ	j	]	}	HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI
						PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS
000	1,550	1 ^	4323	'''	334	
1		İ	1		ł	DPPASASRVAGTTGARHHTQLIFVFLVETGFH
		j				MLARDGLKLLTSSDPPASASQSSWDYRREPP
		1	1			RLANFFVFLVETGSRYVAQAGVQWLFTGAIP
					· ·	LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS
	i	•			ł	VGSPKAKEALNMLTWRAEQEGGMQFWVSSE
	İ	l				SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE
	1				[	GIPIMRWLSRQRNSLGGFASTQDTTVALKALS
!	ł	ł	1		ł	EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT
4	<b>,</b>					HNRLLLQTAELADGTANGSV/SISANGFGFAI
<u> </u>	l					CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV
'	ĺ					AVKENKDDLNHVDLNVCTSFSGPGRSGMAL .
-	1					MEVNLLSGFMVPSEAISLSETVKKVEYDHGK
						LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA
			1 1		<b>[</b>	SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD
			1			VORLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLLSSPWGRAVPC
002	1932		4540	1903	293	
						VSGGLPKPANITFLSINMKNVLQWTPPEGLQG
•		İ				VKVTYTVQYFIYGQKKWLNKSECRNINRTYC
	i					DLSAETSDYEHQYYAKVKAIWGTKCSKWAE
	1				• 0	SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP
	ŀ	ļ			ļ	EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT
						KSNRTWSQCVTNHTLVLTW\LEPNTLYCVHV
						ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK
			ŀ			IIFWYVLPISITVFLFSVMGYSIYRYIHVG\KEK
		ĺ			,	
						HP\ANLILIYG\NEFDKRFFVPA\EKIV\NFI\TL
		'				NISODSKISHQDMSLLGKSSDVSSLNDPQPSG
						NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\
					l	EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI
						CAGPEEQELSLQEEVSTQGTLLESQAALAVL
				İ		GPQTLQYSYTPQLQDLDPLAQEHTDSEEGPEE
				1		EPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCE
						PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY
						LMQFMEEWGLYVQMEN
603	.1953	<u> </u>	4542	-	600	
003	.1700	Α	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL
					i	GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI
						LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL
						LHYAAETGNGEIVKYILDHGPSELLDMADSE
						TGETALHKAACQRNRAVCQLLVDAGASLRK\
				-		TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN
						YKVIGHEDLETAV
604	1954		4540	<del></del> -	030	
304	1734	Α	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN
						QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW
					ĺ	STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG
						QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS
- 1					I	OGOOTOSSAYGSSYTYPPSSLGGTVVDGOPG
			0	, ,		QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEOGMVGLKIGDVSSSA
	: 1					FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA
	: 1					FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS
	:					FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ
605	1955	A .	4553	2	2304	ILQEKRNCLLMQLEEATRLTSYLQSQLKSLC ASTLTVSSGSSRGSLASSRGSLASSRGSLSSVS FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE GPGLGALDRLRAHASAMGDEDLPGMAALQP HGVPGDGEGPHERGPPPASAPVGGTVTLRED SAKRLERRARRISACLSDYSLASDSGVFEPLT KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV HSSALTLKSLQLYVCSVTPQLEELLGIAQIN LADYDSLSEMQLRWHSVQVFTSUNHQGRGR LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR TTAQLQAVERELAEERAKLEYTEEEVLEMER KEEQAEAISERSWQADSVDSGCSNCTQTSPPY PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL KVDKETNTEDLFLEEAASLVKERPSRRARGSP FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL MARTSLDLELDLQASRTRQRQLNEELCALRE LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR EAERQTRQTKLDYRHEQAAEKMLKKASKEI YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL PADDV
606	1956	A	4555	3429	776	PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP VLLLQDSSGDYSLAHVREMACSIVDQKFPEC GFYGMYDKILLFRHDPTSENILQLVKAASDIQ EGDLIEVVLSASATFEDFQIRPHALFVHSYRA PAFCDHCGEMLWGLVRQGLKCEGCGLNYH KRCAFKIPNNCSGVRRRLSNVSLTGVSTIRT SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA PKVPNNCLGEVTINGDLLSPGAESDVVMEEG SDDNDSERNSGLMDDMEEAMVQDAEMAMA ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP LMRVVQSVKHTKRKSSTVMKEGWMVHYTS KDTLRKHHYWRLDSKCITLFQNDTGSRYYKE IPLSEILSLEPVKTSALIPNGANPHCFEITTANV VYYVGENVVNPSSPSPNNSVLTSGVGADVAR MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR NEVAILQNLHHPGVVNLECMFETPERVFVVM EKLHGDMLEMILSSEKGRLPEHITKFLITQILV ALRHLHFKNIVHCDLKPENVLLASADPFPQV KLCDFGFARIIGEKSFRRSVGTPAYLAPEVL RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL RELECKIGERYITHESDDLRWEKYAGEQGLQ YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1957	A	4563		4499	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI QPGAFRRLRNLNTLLLNNNQIKRIPSGAFEDL ENLKYLYLYKNEIQSIDRQAFKGLASLEQLYL

CCC TE	I CEO TO	1 1 4 - 1	050	T Dec 20 1 1 1 1	16.0	T
SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	11100	in NO:	beginning nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
200		ŀ	717	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		i	peptide	3040000	/=possible nucleotide deletion \=possible
				sequence		nucleotide insertion
	$\vdash$		<del>                                     </del>		<del></del>	HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
	1	1				VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
		1	1			DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
	1	1	İ		ĺ	TPEELNCERPRITSEPQDADVTSGNTVYFTCR
		1				AEGNPKPEIIWLRNNNELSMKTDSRLNLLDD
	ľ	1			1	GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
	i	1			İ	EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
	1	ļ				ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS
		ļ	İ	T .		GGLYIQNVVQGDSGEYACSATNNIDSVHATA
	1	ì		1		FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
	Į.	l	]			GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
			1			SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ
		1	1			PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
	}	] [	}			EPAITWNKDGVQVTESGKFHISPEGFLTINDV
			1			GPADAGRYECVARNTIGSASVSMVLSVNVPD
			Ì			VSRNGDPFVATSIVEALATVDRAINSTRTHLF
	1	İ				DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
						ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
	ŀ					PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY
	1					RTHDGTCNNLQHPMWGASLTAFERLLKSVY
	1					ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
	ļ	l				GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
						VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
						DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
						VYPREQINQLTSYIDASNVYGSTEHEARSIRD
						LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
						RDENESPIPCFLAGDHRANEQLGLTSMHTLW
	) .		]			FREHNRIATELLKLNPHWDGDTTYYETRKIVG
		•				AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
			1 1			PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD
						ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
	1					RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
						VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
						AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID LFPALVVEDLVPGSRLGPTLMCLLSTOFKRLR
						DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
	1			Į.		NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
				İ		LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
						FSYQEDKPTKKTRPRKIPSVGROGEHLSNSTS
					1	A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR
						TOIKKLESR\LSTTECVDAGGESHANNTKWK
				1	ĺ	KDACTICECKDGQVTCFVEACPPATCAVPVNI
						PGACCPVCLQKRAEEKP
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
					1	IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
		1		١ ٠ ١		GLNQCMSGIINHEAYHEVPYTTSFTLAKOLSF
				l		YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ
i				ŀ	ł	LGVGNYKKRLGINLLGGPLGGKQVIRVSCGD
				ļ		EFTIAATDDNHIFAWGNGGNGRLAMTPTERP
			ļ <b>l</b>	ļ	ĺ	HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
			ł	l	1	VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
						GGPDAW
609	1959	Α .	4567	1	412	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
•						PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
			.			HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
			1		1	RARPRINLRNVTYSFAVTYCLNYISLAMSSTL
	10.5					KLSFHVLSGS
610	1960	A	4570	697	467	ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
	1		· ·		1	GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N
			1			LVIRPPRPPKVLGLQA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ĺ	ŀ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ	ļ			peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
						WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
	•				1	LCAATAVLLSAQGGPVQSKSPRFASWDEMN
1		ſ		[		VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
1		ļ				LSACGSACOGTEGSTDLPLAPESRVDPEVLHS
1		İ			į	LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
1		ł			1	RIQHLOSOFGLLDHKHLDHEVAKPARRKRLP
					Ì	EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
						QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIOR
1.						RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
1						EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
1			1		1	VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG
ł						LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
]						GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
612	1062		1676	160		WRGRYYPLQATTMLIQPMAAEAAS
612	1962	A	4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
(13	1062	•	4604		201	GSPASASPVAGITGTRHHRTRG
613	1963	Α	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
1						SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
						HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
						LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	Α	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
†						GGLFCAWVGTILLVVAMATDHWMQYRLSGS
						FAHQGLWRYCLGNKCYLQTDSIAYWNATRA
						FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
i						GIFYMCAYRVHECRRLSTPR
615	1965	Α	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
<b>i</b> i						ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
!						PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
1						ERRPLPKNARPSPWVPALADEWNTLHQEVTT
						TRLPAGSQEPVKD
616	1966	Α	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
i i		-				SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
						AGLELPTSGDPPALASQSAGITGVTTVPSGPG
617	1967	В	4595	84	478	XRHGLREPLLERRCAAASSFQHSSSLGRELPY
					.,,	DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
	1					REKQLQELQQQEEEERQRQQRREERRQQNL
						RARSREHPVVGHPDPALPPSGVNCSGCGAEL
			1			HCQDAR*
618	1968	A	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
113	.,,,		,		1100	SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
ŀ	I					MVGNETTYEDGHGSRKNITDLVEGAKKANG
	ŀ	ł	ł	1		
					ļ	VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA MSLLSIILLHLLAGIMGWVMIIMEISELGYRIF
			}	i	I	HCYMEYSRLRGEAGSDVSLVDLGFOTDFRV
	l	l	ŀ			
	I	1	l			YLHLRQTWLAFMILSILEVIIILLLIFLRKRILI
	l		ļ	1	ł	AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
	1					
		J	ł			CNPETFPSSNESRQCPNARCQFAFYGGESGYH
			ł			RALLGLQIFNAFMFFWLANFVLALGQVTLAG
	ĺ		[	4	ſ	AFASYYWALRKPDDLPAFPLFSAFGRALRYH
	l	1	1	l		TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
	j	l		ŀ	ļ	KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
	j	l		<b>!</b>		IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
	1	ļ	-	1		TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
	1	1				APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC
	į		ł	ľ		VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
	10.50					LNKTNKKAAES
619	1969	Α	4601	2	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY

NO: of nucleotide cotid	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucle odide sequence   peptide continue sequence   peptide sequence						1	D=Aspartic Acid. E=Glutamic Acid.
Contraction   Contraction	nucl-	peptide	ł	in		location	
Seq.   uence   9949   correspondi   914   for first st mino acid residue of peptide   sequence	eotide	seq-	1	USSN	location		I=Isoleucine K=I vsine I =I eucine
gl of first amino acid residue of peptide residue of peptide of	seq-	uence		09/496	correspondi		M=Methionine N=A sparagine P=Proline
	uence		1				O=Glutamine R=Arginine S=Serine
Pepidde   Sequence		1					T=Threonine V=Valine W=Truntonhan
Poptide	1	1	ĺ	İ			V=Tyrosine Y=I labroum *=Ston codon
Sequence	1		1			3040000	/=nessible nucleatide deletion \=nessible
GLAYGHRYFESNHEDWVDLQGWYTGNG   GLYLTDPQHISVDQKYTTHYGRGFFYFFN   NOHVECNEICHRLSLTRPMEKPCKS    MERLWGLFQRAQQLSPRSSQTVYQRVEGFR   KGHLEEEEDGEGAPTLAHFCPMELRGFPP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPWAAAGRAAPYLVLTALGFP   LGSRPRQFPALPAGSAGMAALTQDIRAALS   RQKLDHVWTDTHYVGLQFPPAHPNTLW   DEAGKVGGQLPLEDPDVYCFYSAIGNVTGEL   VYAHYGRPEDLQDLRARGVDPVGRLLLYPEPADFSQ   GVISFAQKVTHAQDFGAQGVLYPEPADFSQ   GVISFAQKVTHAQDFGAQGVLYPEPADFSQ   GVISFAQKVTHAQDFGAQGVLYPEPADFSQ   GVISFAQKVTHAQDFGAQGVLYPEPADFSQ   DPFKPSLSSQQAVGHVHUGTGDPYTPGPBP   NOT AN AN AN AN AN AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AN AN AN AN SERVERY AND AN AN AN SERVERY AND AN AN AN SERVERY AND AN AN AN SERVERY AND AN AN AN SERVERY AND AN AN AN SERVERY AN AN AN SERVERY AND AN AN SERVERY AND AN AN AN SERVERY AND AND AN AN SERVERY AND AN AN AN AN AN AN AN AN AN AN AN AN AN	ľ	}	ł	l			
GLIVITIPOGHISVDOKVFTINFGKRGFYFFN			_		Boquesion	<del> </del>	
NOHVECNEICHRISTRYMEKPCKS	Į		Į.	1	ŀ		
1970   A   4666   1   2415			I		į		· · · · · · · · · · · · · · · · · · ·
KGHLEEEEIOGEEGAETLAHFCPMELRGPPE LOSRPRQPNLIPWAAAGRRAAPYLVTALLIE TGAFLLGYVAPRGSCQACGDSVLVVSEDVN YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL EDTTRQTSLERWAGSAGMALTQDIRAALS ROKLDHVWTDTHYVGQFPDPAHPNTLHW DEAGKVEGOLDELDPDVYSAGNOTTGEL VYAHYGRPEDLQDLRARGVDPVGRLLLVRV GVISFAQKVTNAQDFGQQGVLIYPEPADFSQ DPFKPSLSSQQAVYGHVHLGTGDPYTPGFFSI NQTGFPPVASSGLFSIPAQPISADIASRLIRKL KGPVAPQEWQGSLLGSPTYHLGFGRELRLVV NIRTSTPINNIFGGEGRSEFDHVVVIGAQRD WGFGAAKSAVGTALLELTVATTSSMVSNGFR PRSLLFISWDGGDFGSVGSTEWLEGYLLVRV HLKAVVVVSLDNAVLGDBKATSFILTSSLIKAL GVIRPLPMDSSAYISFTAFVGVPAVEFSME DDQAYFFLHTKEDTYSENLKVLQGRLPAVA QAVAQLAQQLIRLSHDRLIPLDFGRYGDVV LRHIGHNLEFSGDLKARGLTLQWYYSARGDY RAAEKLRQEIYSSEERDER.TRYMYNKIMEV EFVFLSQVVSPADSFFRHIFMGRGDHTIGALL DHIAILASNSSGTPGATSSTGFQESSFRRQLA ALLTWDACKGANALSGDWIDMINF GSSIVVFFTMRIGTSISAASIS-FLGLGAQPTPT EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLADALPFKKG GSSIVVFFTMRIGTSISAASIS-FLGLGAQPTPT EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLADLDFKKG SSIVVFFTMRIGTSISAASIS-FLGLGAQPTPT EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLADLDFKKG CTLACRGKNMEVRLIFLSGCLAVAVVWAVF RNEDRWAWRQDLIGAFCLINLKTLKLPFK SCVLLIGLILLYDVFFVTFTRIGGSIMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPKLIYFSVMSVCLMFVSRLGFGDIVFGL LIAYGRRFDVGTGSSYTYVSVTVAYAIGML TFVVLGLMKKQQPALLYLVPCTLITTACQFV AWETVREMKKFWERVTS CTLACRGKNMEVERLIFLSGGSMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPLMKKFWERVTS CTLACRGKNMEVERLIFLSGGSMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPLMKKRYBENTS DHYTYLEMKKFWERVTS CTLACRGKNMEVERLIFLSGGSMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPLMKKRYBENTS CTLACRGKNMEVRLIFLSGGSMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPLMKKRYBENTS CTLACRGKNMEVRLIFLSGGSMVSL AGFFGNNEKNDGNLVEATGQFSAPJEKLPV VIRVPLMMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVLLYVCAFYNBAAAMNWRLFSKY QYFDSRGMFISIVFSAPLIVMGTWMVWW KTLNVMTDLKNAQERKEKKRKREGUGAA AAWSLRSSRFSPAAPSAAVCVWASSFQLTHG LKNRCFI LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQFESANDFEPLCAVCGGAH	620	1970	A	1606	<del></del>	2415	
LGSRPQPNLPWAAGRRAAPY.UTALLIE TGAFLLGYVAPRGSCQACGDSVLVVSEDVN YEPDLDFHGGRLYWSDLQAMFLQFLGEGRL EDTIRQTSLRERVAGSAGMAALTQDRAALS RQKLDIHVWTDTHYVGLQFPDPAHPNTLHWV DEAGKVGEQLPLEDPDVYCPYSAGNTGEL VYAHYGRPEDLQDLRARGVDPVGRILLVRV GVISFAQKVTNAQDFGAQGVLYPEPDLPSDFSQ DPPKPSLSSQQAVYGHVLGTGDPYTFGFFSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGTGDPYTFGFFSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGTGDPYTFGFFSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGTGDPYTFGFFSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGTGDPYTFGFSSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGTGDPYLTVIGAQRDA WGFGAAKSAVGTAILLELVRTFSSWSNOGFF PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL HLKAVYYSLDNAVLGDDKFHAKTSPLLTSSI IESSYLKQVOSPANDSVGSTEWLEGYLSVL HLKAVYYSLDNAVLGDDKFHAKTSPLLTSSI IESSYLKQVOSPANDSFRHIPMGRGDHTLGALL DPHATLLRSNSGTJGGATSPTPCVSFTRAPSWD AEVIRPLMSSSGTJGGATSFQDVSTRIMW EFYFLSQVVSPADSPFRHIPMGRGDHTLGALL DHIRLLRSNSGTJGGATSFQDVSTRIMBPGT GSSTVVFFTMRIGTSIISAASDTVLLRHILPGT GSSTVVFFTMRIGTSIISAASDTVLLRHILPGT GSSTVVFFTMRIGTSIISAASDTVLLRHILPGT GSSTVVFFTMRIGTSIISAASDTVLLRHILPGT GSSTVVFFTMRIGTSIISAASDTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAGADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAADMTVLLRHILPGT GSSTVVFFTMRIGTSIISAASSTVDGAADMTVLLRHILPGT GTACRKNMEVRLIFLSGLCLAVAVVWAVF RNEDRWAWLQDDLGJAFCLNLIKLLPHFYG CTLACRGKNMEVRLIFLSGLCLAVAVVWAVF RNEDRWAWLQDDLGJAFCLNLIKLLPHFYG CTLACRGKNMEVRLIFLSGLCLAVAVVWAVF RNEDRWAWLQDDLGJAFCLNLIKLPHFYG CTLACRGKNMEVRLIFLSGLCLAVAVVWAVF RNEDRWAWLQDDLGJAFCLNLIKLPHFYG GGGCCGSGLJAFTMPASTSVSTPVSIGAGDITVPGL LLAYCRFDVQGTSSTVTYSSVTVAJAGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVVLGLMKKGQPALLYLVATAJGML TTVLGLMKGQPALLYLVATAJGML TTVVLGLMKAGAAMMVRLPSKY QYFDSRGMF	020	1370	^	4000	'	2415	MEKL WGLFQKAQQLSPKSSQTVYQKVEGPR
TGAFLGYYAFRGSCQACGDSVLVYSEDVN YEPDLDFHQGRLYWSEDVAM YEPDLDFHQGRLYWSEDVAM YEPDLDFHQGRLYWSEDVAM YEPDLDFHQGRLYWSEDVAM RQKLDHVWTDTHYVGLQFFDPAHPNLHWV DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL VYAHYGRPEDLQDLRARGVDPVGPLSADFSQ GVISFAQKVTNAQDFAGQVLIYPEPADFSQ GVISFAQKVTNAQDFAGQVLIYPEPADFSQ DPPKPSLSSQQAVYGHVHLGTGDPYTFGFFSSI NQTOFPPVASSGLPSPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPHQFGFRIRLVVN NIHISTIPINNIFGCEGRSEPDHYVIGQQRDA WGGGAAKSAVGTAILELVRTFSSMVSNGFR PRSILFISWDGGDFGSVGSTEWLEGYLSVL HLKAVVYVSLDNAV.LGGGFASVSTFWLEGYLSVL HLKAVVYVSLDNAV.LGGHFAKTSPLLTSI IESVLKQVDSPNHSGQTLYEQVVFTNPSWDI AEVIRPLPMDSSAVSFTAFVGVFAVEFSFME DDQAYPPLHTKEDTYENHKVLQGRLPAVA QAVAQLAGQLLIRLSHDRLLPLDFGFYGDVV LRHIGNLINFSGDLKARGTLAWVYSARGDV IRAGEKLRQEIYSSEERDERLTRMYVVRIMRV EFYFLSQYVSPADSPFRHIMGGGDHTLGAUL HHIGNLINFSGDLKARGTLAWVYSARGDV IRAGEKLRQEIYSSEERDERLTRMYVVRIMRV EFYFLSQYVSPADSPFRHIMGGGDHTLGAUL HHIGNLINFSGDLKARGJANDHVINNFF ISRVDDFVGSGIANVIUAVISPIPAFARLVRGT NILVLKQQTFIESARSIGASDMTVLRHILIPGT GSSIVVFFTHRIGTSISAASLSFLGLGAQPFTP EWGAMLNEARAMVIAPHVAVFPALAIFLTV LAFNLGDGLRDALDPKIRG  621 1972 A 4614 2 820 LVYVMAHFCIASAMSLYNCIAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF RNEDRWAWMLQDLGJAGLALIHKIPYGGV CTIACRGKNMEVRLIFLSGLCAVAVWAVF AWSLRYSSMYSSSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ GGCCESGLPNTMPSARSVSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ GGCCESGLPNTMPSARSVSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ GGGCESGLPNTMPSARSVSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ GGGCESGLPNTMPSARSVSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ GGGCESGLPNTMPSARSVSFPVSIPAVLTQT DWTEFWLMGLATTHALCVLLTCLSSRSYRLQ AWSLRPSRPPSAAPSAANCVAWASSQLTHG LKNRCFI							KUHLEEEEEDGEEGAETLAHFCPMELRGPEP
SPENDLOFHGGRLYWSDLQAMFLCPIGLGGRL   EDTIRQTSLERNAGSAMALTQDIRAALS   RQKLDHVWTDTHYYGLQFPDPAHFNTLHW  DEAGKVGEQLPLEDPDVYCPYAGNATGTORAALS   RQKLDHVWTDTHYYGLQFPDPAHFNTLHW  DEAGKVGEQLPLEDPDVYCPYAGNATGTORAGVLLYRVGVGAGQSLLYPEADPSQ   DPPKPSLSSQQAVYGHVLGTGDPYTTGFFSI   VYAHYGRPEDLQDLARAGVDPYGRLLLVRV   GVISFAQKVTHAQDFGAGQSLLYPEADPSQ   DPPKPSLSSQQAVYGHVLGTGDPYTTGFFSI   NQTOFPPVASSGLSPAQPISADIASRLLKRL   KGPVAPQEWQGSLLGSPHQFGRELLVVN   NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRD/   WGFGAAKSAVGTAILLELVRTFSSMYNGFF   PRRSLLFISWDGQDFGSVGSTEWLEGYLSVL   HLKAVVYVSLDHAVLGDDKFHAKTSPLLTSL   EESVLKQVDSPNHSGGTLVVVYSARGDV   AZVIRPLPMDSSAYSFTAFVGVPAVEFSMW   DDQAYPPHTKEDTYSHLKVLQGRLPAVA   QAVQLAGQLLIRLSHDRILPLDFGRYGDVV   LRHIGNINEFSGDLKARGITLQWVYSARGDV   RAAEKLRQETYSSEERDFRITMYNVRIMRV   EFYFLSQYVSPADSPFRHIFMGGGDFTLGALL   DHIRLILRSNSGTPGATSFRQLVAG   ALLIYWDACKGAANALSGDVWHIDNNF   EFYFLSQYVSPADSPFRHIFMGFGDHTLGALL   DHIRLILRSNSGTPGATSFRQLVAG   ALLIYWDACKGAANALSGDVWHIDNNF   GSSIVVFFTMRIGTSIISAASLSFLGGAQPFTP   EWGAMLNEARADMVILAPHVAVFPALAIFLTV   LAFNILGDGLRDAJDPKIKG   SSIVVFFTMRIGTSIISAASLSFLGGAQPFTP   EWGAMLNEARADMVILAPHVAVFPALAIFLTV   LAFNILGDGLRDAJDPKIKG   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLILPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLIPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLIPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLIPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDLGJAFCLNLIKLIKLIPPGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDRWAWILQDL	ĺ	1	ί.	1		l	LUSKPKQPNLIPWAAAGRKAAPYLVLTALLIF
EDTRGTSLRERVAGSAGMAALTQDRAALS RQKLDHUWTDTHYVGLQFPDAHPNTLHW DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL VYAHYGRPEDLQDLRARGVDPVGRLLLVRV GVISFAQKVTMAQDFAQQVLIYPEPADFSQ DPFKPSLSSQQAVYGHVHLGTGDPYTTGFPSI NQTGFPVASSGLPSISAQPISADIASALIKKL KGPVAPQEWQGSLLGSPYHLGPGPRIRLVVV NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRD/ WGPGAAKSAVGTAILLELVRTFSSMYSNGFR PRRSLLFISWDGGDFGSVGSTEWLEGVYLSV HLKAVVVSLDNAVLGDDKFHAKTSPLLTSI IESVLKQVDSPHSGQTLYEQVVFTNIPSWD) AEVERPLEMDSSA YAVVYSLDNAVLGDDKFHAKTSPLLTSI IESVLKQVDSPHSGQTLYEQVVFTNIPSWD) AEVERPLEMDSSA YAFTAFVGVPAVEFSFME DDQVAYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLRLSHDRLLPLDFGRYGDVV LARHIGNINFSGDLKAGTLTQWVYSARGDY IRAAEKLRQEIYSSEREDERLTRMYNVRINMEV EFYFLSQVYSPADSFRIEITMGRGGHTU-GALL DHLRLRSNSGTPGATSSTGFQVESRFRRQL\ ALLYTWDACKGAANALSGDVMIDNINF ISRYDDFVGSGJANVILAVHFSIPAFARLVRGI NITLVLKQQTIFIESARSIGASDMTVLLRHILPGI GSSIVVFTMRIGTSIBASLSFLGLGAQPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG CTIACRGRNMEVRLIFLSGLCIAVAVVWAVF REDERWAWILQDILGCALNLKTKLRIPFK SCVILLGLLLLYDVFFVFTIFTIKNGESIMVEL AAGPFGNNEKNDGNLVCAALHHKIPYGQ CTIACRGRNMEVRLIFLSGLCIAVAVVWAVF REDERWAWILQDILGCALNLKTKLRIPFK SCVILLGLLLLYDVFFVFTIFTIKNGESIMVEL AAGPFGNNEKNDGNLVCAALHHKIPYGQ CTIACRGRNMEVRLIFLSGLCIAVAVVWAVF REDERWAWILQDILGCALNLKTKLRIPFK SCVILLGLLLLYDVFFVFTIFTIKNGESIMVEL AAGPFGNNEKNDGNLVCAACH LIAYCRRPDVQTGSSYTYYVSVIVAYAIGMIL TFVVLGJMKKGQPALLVLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSSSYRLQ QFGRSGMFISIVFSAPLLVNAMILVNMVVW KTLNVMTDLKNAQERRKEKKRRRED*GAA AAWSLRPSRPPSAAPSAVCVAWASSFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1				1	IGAFLLGYVAPRGSCQACGDSVLVVSEDVN
RQKLDHVWTDTHYVGLQFPDPAHPNTLHW DEAGKVGEQILELDPYCTYSAGNYTGEL VYAHYGRPEDLQDLRARGVDPVGRLLLVRV GVISFAQKVTMAQDFGAQGVLIYPEPADFSQ DPKPSLSSQQAVYDRLGTGDPYTFGFFSI NQTGFPPVASSGLSSQAVYGHUGTGDPYTFGFFSI NQTGFPPVASSGLSSGAVGPHAGTGDPYTFGFFSI NQTGFPPVASSGLSSGAVGPHAGTGDPYTFGFFSI NQTGFPPVASSGLSSGAVGPHAGTGDFYTFGFFSI NQTGFPPVASSGLSSGAVGFBGAGPHAUVVIGAQRDA WGPGAAKSAVGTALLELVRIFTSSMYSNGFR PRSILIFISWDGGDFGSVGSTEWLEGYLSVI HILKAVVYVSLDNAVLGDDKFHAKTSPLLTSI ESVLKQVDSPNHSGGTLYEQVVYTNIPSWD) AEVIRPLPMDSSAYSFTAFVGVPAVEFSFME DDQAYFFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAQQLIRLSHDRLIPLDFGFKGDIVV LRHIGGILNEFSGLKARGLTLQWVYSARGDY RAAEKLRQETYSSEBELTRMYNVRIMRV EFYFLSQVVSPADSPFRHIPMGRODHTLGALL DHLRLRSNSSGTPGATSSTGPGESTRRQLY ALLTITVDACKGAANALSGDVMIDDNIP  621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSPAFARLVRG, MILVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIBAASLSFLGLGAQPFTP EWGAMLNEARADMVIAPHVAVPPALAIFLTV LAFINLIGGILRDALPKIKG  622 1972 A 4614 2 820 LVYVMAIFCLGSAMSLYNCLAALIHKIPYGG CTIACRGKNMEVRLIFLSGLCLAVVVWAVF RNEDRWAWILQDILGJACCHNLIKTLKLPNFK SCVILLGLLLLYDVFFYFITPTITKNGESIMVEL AAGPFGNNEENDGOLVEATGQPSAPHEKLPV VIRVPKLITYSVMSVCLMPVSILGFGDIIVPGL LLAYCRFDVQTGSSYTYTVSVIVTAYAIGMIL TVVLGLMKKGQPALLVLVPCTLITA/QFV AWETVREMKKFWERVITS  623 1973 A 4619 17 691 TLVSVVEPVRRADLITREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSSSYBLQ GGHTLCLVLLVYCAFBAAMNIWLFSKY QVFDSRGMTISIVFSAPLLVNAMILVNMVW WXTLNVMTDLKNAQERRKKRRRED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	l	ł	i	}		ł	YEPDLDFHQGRLYWSDLQAMFLQFLGEGRL
DEAGKVGEQL/PLEDPDYVCYSAIGNVTGEL VYAHYGRPEDLQDILARGVOPPYGRLUVRV GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSI NQTGFPYVASSGLPSIPAQPISADIASRLIKKL KGPVAPQEWQGSLLGSPYHLGPGPRIRLUVVX NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRD/ WGPGAAKSAVGTAILLELVRTFSSMYSNGFR PRRSLLFISWDGGDFGSVGSTEWLEGYLSVIL HLKAVVYSLDNAVLGDDKFHAKTSPLTSI IESVLKQVDSPNHSGQTLYEQVVFTNNPSWDI AEVIRPLEMIDSSAYSFTAFVGPVAEFSFME DDQVAYPFHTHEDTYENLHKVLQGRLPAVA QAVQLAGQLIRLSHDRILPLDFGRYGDVV LRHIGGNLFSSGDLKAGLTLQWVYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLIRSNSSGTPGAGTSTGPQESFRRQLIA ALLTWDACKGAANALSGDVWMIDNNF 621 1971 A 4610 793 334 ISRVDDFVGSGIANVILAVAIFSIPAFARLVRGI GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTPF EWGAMLNEARADNAPHYAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMAIFCIASAMSLYNCLAALIHKIPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTPT EWGAMLNEARADNAPHYAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMAIFCIASAMSLYNCLAALIHKIPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTPT EWGAMLNEARADNAPHYAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 2 820 LVYVMAIFCIASAMSLYNCLAALIHKIPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPTPT EWGAMLNEARADNAPHYAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG 622 1972 A 4614 1 17 691 TVVVWERMKKWERVIS 6CVILLGLLLLYDVFFVFTIFTITKNGESIMVEL AAGPFGNNENDGNLVBATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIVYGL LLAYCRRFDVQTGSSVTYSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYSTQL GGGCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYSTQL GGGCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYSTQL GGGCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYSTQL GGGCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYSTQL GGGCESGLPNTMPSAFSAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGGAH	İ			1		İ	EDTIRQTSLRERVAGSAGMAALTQDIRAALS
SYAHYGREEDLQDLRARGVDPYGRLLLVRY   GVISFAQKVTNAQDFAQFSQQVLYPEPADFSQ    DPFKPSLSSQQAVYGHVHLGTGDPYTFGFPSI    NQTQFPPVASSGLPSIPAQPISDADIASRLLRKL    KGPVAPQEWQGSLLGSYPH.GPGFRILVVN    NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA    WGPGAAKSAVGTAILLELVRTFSSMVSNGFR    PRSLLFISWDGGDFGSVGSTEWLEGYLSVL    HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSI    ESVLKQVDSPHISGQTLYQVYFTNBYSWD    ACVIRPLEMDDSAYSFTAFVGVPAVEFSEME    DDQ\AYPFHTKEDTYENLHKVLQRVLPAVA    QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV    LRHIGNLNEFSGDLKARGLTLQWYYSARGDY    RAAEKLRQEIYSSERDERLTRMYNYRIMRV    EFYFLSQYSPADSFFRIIFMGRGDHTLGAIL    DHLRLLRSNSSGTPGATSSTGPQESSFRRQLV    LRHIGNLNEFSGDLKARGLTLQWYYSARGDY    LRALKLQEIYSSERDERLTRMYNYRIMRV    EFYFLSQYSPADSFFRIIFMGRGDHTLGAIL    DHLRLLRSNSSGTPGATSSTGPQESSFRRQLV    ALAINLEQUENSASIGASDMTVLLRHILPGT    GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP    EWGAMLNEARADMVIAPHVAVPPALAIFLTV    LAFINLEDGLRDALDFKIKG    622	1		ļ			1	RQKLDHVWTDTHYVGLQFPDPAHPNTLHWV
GYISFAQKVTNAQDFGAQQVLIYPEPADFSQ DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSI NQTQFPPVASSGQAVYGHVHLGTGDPYTPGFPSI NQTQFPPVASSGQAVYGHVHLGTGDPYTPGFPSI NQTQFPPVASSGLGSPYHLGPGPRIRLVVN NHRTSTPNNIFGGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRTFSSMVSNGFR PRRSLLFSWDGGDFGSVGSTEWLEGYLSVL HLKAVVYSLDHAVLGDDKFHAKTSPLLTSI IESVLKQVDSPNHSQQTLYEQVVFTNPSWDN AEVIRLPMDSSAYYSFTAPKGVPAVEFSFME DDQAYPFLHTKEDTYENLHKVLQGRLPAV QAVAQLAQGLIRLSHDRLIPLDFGRYGDVV LRHIGNINEFSGLAKAGRLTLQWVYSARGDY RRAÆKLRQETYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFFHIFMGRGDHTIGALL DHLRLIRSNSSGTPGATSSTGFQESRFRRQL) ALLTWDACKGAANALSGDWNIDNNF 1SRVDDFVGSGIANVILAVAIFSIPAFARLVRGI NTLVLKQQTFIESARSIGASDMYVLLRHILPGI GSSIVVFFTMRIGTSIISAASLSFLGLGAQPFTP EWGAMINEARADMVIAPHVAVPALAIFILTV LAFNLLGDGIRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALHKIPYGQ CTIACRGKNMEVILIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVTIPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSLGFGDIIVYGL LIAYCRFFDVQTGSSYTYYSVTVAYAGMUL TFVVLGUMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAFSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAFYINEAAAMNWRLFSKY QYPDSRGMFISIVFSAPLLVAMMIVVWWW KTLNVMTDLKNAQERKKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNOPESANDPEPLCAVCGOAH							DEAGKVGEQLPLEDPDVYCPYSAIGNVTGEL
DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSI NQTQFPPVASSGLPSIPAQPISADIASRLLRKL KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVY NHRTSTPINNIFGCEGRSEPDHYVVIGAQRDA WGPGAAKSAVGTAILLELVRITSSMYSNGFR PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL HLKAVYVYSLDNAVLGDDKFHAKTSPLLTSL IESVLKQVDSPNHSQTILYEQVYFNPSWD) AEVIRPLPMDSSAYSFTARVGVPAVEFSFME DDQVAYPFLHTKEDTYENLHKVLQCRLBAVA QAVAQLAGQLLRISHDRILPLDFGRYGDVV LRHIGNI.NEFSGDLKARGLTLQWVYSARGDV IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYTLSQYVSPADSPFRHIFMGRGDHTLGALL DHLRLB.SNSSGTFGATSSTGFQUSSFRRQLA ALVITVDACKGAANALSGDVWIDDINF  621 1971 A 4610 793 334 ISRVDDFVGSGIANVILAVAIFSIFAFARLVRGI NILVLKQQIFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIFSIISAASISSMTVLLRHILPGT GSSIVVFFTMRIFSIISAASISSMTVLLRHILPGT GSSIVVFFTMRIFSIISAASISSMTVLLRHILPGT GSSVVFFTMRIFSIISAASISSMTVLLRHILPGT GCSVVFFTMRIFSIISAASISVNCLAALHKRIPYGQ CTTACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDLGJAFCLNIKITLKIPKK SCVILLGLLLLYDVFFVFITPFITNGGESMYEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVTVAYAIGMIL TTVVLGUMKKGOPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTOT DWTTEPWLMGLATHALCVLLTCLSSSSYRLQ UFFDLATIONAGERKKEKKRRKED/GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNOPESANDPEPLCAVCGOAH	ĺ		l				VYAHYGRPEDLQDLRARGVDPVGRLLLVRV
NOTOFPPVASSGLPSIPAQPISADIASRLLIKL   KOPVAPOEWOSLLGSPYHLGPGPRLLVVN   NHRISTPINNIFGCEGRSEPDHYVVIGAQRDA   WOFGAAKSAVGTAILLEURITSSMYSNOFF   PRESILFISWDGOFGSVOSTEWLEGYLSVL   HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSI   IESVLKQVDSPNHSQGTLVEQVVFINSWDJ   AEVIRPLPMDSSAYSFTAFVGVPAVEFSFME   DDQIAYPFLHTKEDTYENLHKVLQGRLPAVA   QAVAQLAGQLIRLSHDRLLPLDFGRYGDVV   LRHIGNINEFSGDLKARGLTLQWYVSARGDY   IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV   EFYFLSQVVSPSDEPRIHMGRGDHTLGALL   DHLRLLRSNSSGTPGATSSTGFQESRFRRQL   ALLITWDACKGAANALSGDVWNIDNNF   GSSIVVDFVGSGANVILVALRSIPAFARLVRG   NTLVLKQQTFESARSIGASDMTVLLRHILPGT   GSSIVFTMRIGTSIISAASLSFLGLGAQPFTP   EWGAMLNEARADMVIAPHVAVFPALAIFLTV   LAFRILLGDGLRDALDPKIKG   LAYVMALFCIASAMSLYNCLAALIHKIPYGQ   CTHACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDR WAWILQDILGJAFCLNLIKTLKPNFK   SCVILLGLLLYDVFFVFTIPFTINGGESIMVEL   AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV   VIRVPKLIYTSVMSVCLMPVSLGFGDIDVPGL   LIAYCRRFDVGTGSSTYTYVSVTVAYAIGMIL   TFVVLGLMKKGQPALLYLVPCTLITA/CQFV   AWETVREMKKFWERTVTS   AWETVREMKKFWERTVTS   GGCCESGLPNTMPSAFSVSFPVSIPAVLTOT   DWTEPWLMGLATFHALCVLLTCLSRSYRLQ   IGHFLCLVILVYCAEYINEAAAMNWRLFSKY   QYPDSRGMFISIVFSAPLLVNAMIIVVMVVW   KTLNVMTDLKAORERKEKKRRKED*GAA   AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG   LKNRCF!	ĺ		[			[	GVISFAQKVTNAQDFGAQGVLIYPEPADFSQ
KGPVAPQEWQGSLLGSPYHLGFGPRIRLVVN   NIRTSTPINNIFGCIEGRSEPDHYVVIGAQRD/    WGPGAAKSAVGTAILLELVRITSSMVSNGFR     PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL     HLKAVVYVSLDNAVLGDDKPHAKTSPILTSI     IESVLKQVDSPNHSGQTLYEQVVFTNVPSWD      AEVIRPLPMDSSAYSFTAFVGVPAVEFSFME     DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA     QAVAQLAGQLIRLSHDRLLPLDFGRYGDVV     LRHIGNINEFSGDLKARGLTLQWVYSARGDY     RAAEKLRQEIYSSEERDERLTRMYNVRIMRV     EFYFLSQVVSPADSPFRHIPMGRGDHTLGALL     DHLRLISNSSGTPGATSSTGFQUESRFRRQLV     ALLITWDACKGAANALSGDVWNIDNNF     GSSIVVFFMRIGTSIISAASLSFLGLGAQPFTP     EWGAMLNEARADMVIAPHVAVFPALAIFLTV     LAFNLLGDGLRDALDPKIKG     GSSIVVFFMRIGTSIISAASLSFLGLGAQPFTP     EWGAMLNEARADMVIAPHVAVFPALAIFLTV     LAFNLLGDGLRDALDPKIKG     CTTACRGKNMEVRLIFLSGLCLAVAVVWAVF     RNEDRWAWILQDLGIAFCLNLIKTLKLPNFK     SCVILLGLLLLYDVFFVFTIPFITKNGESIMVEL     AAGFFGNNEKNDGNLVAATGQPSAPHEKLPV     VIRVPKLLYFSVMSVLVAFYGQPSAPHEKLPV     VIRVPKLLYFSVMSVLVAFYGQPSAPHEKLPV     VIRVPKLLYFSVMSVLVAFYGQPSAPHEKLPV     VIRVPKLLYFSVMSVLVAFYGGPSAPHEKLPV     VIRVPKLLYFSVSFYVSTRAVLTOT     DWTTSPWLMGLATFHALCVLLTLCLSSRSYBLQ     GGCCESGLPNTMPSAFSVSFFVSTRAVLTOT     DWTTSPWLMGLATFHALCVLLTCLSSRSYBLQ     GGFLCLVILVYCAFYINEAAAMNWRLFSKY     QYFDSRGMFISIVFSAPAVSTPVSTRAVLTOT     DWTTSPWLMGLATFHALCVLLTCLSSRSYBLQ     GGFLCLVILVYCAFYINEAAAMNWRLFSKY     QYFDSRGMFISIVFSAPALVOTMWVWW     KTLNVMTDLKNAQERREKKRRKED® GAA     AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG     LKNRCTIL	i		l				DPPKPSLSSQQAVYGHVHLGTGDPYTPGFPSF
NIRTSTPNNIFGCIEGRSEPDHYVVIGAQRDA	1	-{					NQTQFPPVASSGLPSIPAQPISADIASRLLRKL
NIRTSTPNNIFGCIEGRSEPDHYVVIGAQRDA	ĺ		l	]		•	KGPVAPQEWQGSLLGSPYHLGPGPRLRLVVN
WGFGAAKSAVGTALLELVRITSSMVSNGFR PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSI ESVLKQVDSPNHSGOTLYEQVVYFINPSWD) AFVIRPLPMDSSAYSFTAFVGVPAVEFSFME DDQAYPFHTKEDTYENLHKVLQGRLFAVA QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV LRHGINLDFSGDLKARGLTLQWVYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFHHIMGRGDHTLGALL DHLRILRSNSGTPGATSSTGFQESSFRRQL) ALLYWDACKGAANALSGDVWNIDNNF 621 1971 A 4610 793 334 ISRVDDFVGSGIANVILAVAIFSIPAFARLVRGO GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVPFALAIFILTV LAFNLLGGDRAALDPKIKG 622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCLAVAVVWAVF RNEDRWAWMLQDILGJAFCLNLIKTLKLPHFK SCVILLGLLLLYDVFFVFTIPFITKNGESIMVEL AAGPFGNNEINDGINLVEATGQPSAPHEKLPV VIRVPKLIFTSYSWSVCLMPVSILGFGDIVPGL LIAVCRRFDVQTGSSYTYYVSVITVAYAIGMIL TFVVLGAMKKGQPALLYLLVPCTLITIA/CQFV AWETVREMKFWERVTS 623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSSPTVSIPAVLTQT DWTEPPWLMWGLATFHALCVLLTCLSSRSYBLQ IGHFLCLVILVYCAFYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKEEP-GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNOPESANDPEPLCAVCGOAH	l	ł		} }		ľ	NHRTSTPINNIFGCIEGRSEPDHYVVIGAQRDA
HIKAVYYSLDNAVLGDDKFHAKTSPLLTSI IESVLKQVDSPNHSQTILYEQVVFTNPSWD\ AEVTRILPMDSSAYSFTAFVGVPAVEFSFME DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRISHDRILPLDFGRYGDVV LRRIIGNI.NEFSGDLKARGLTLQWYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIPMGRGDHTLGALL DHI.RLLRSSGTPGATSSTOFQ\ESRFRRQL\ ALL\TWDACKGAANALSGDVWNIDNNF  621 1971 A 4610 793 334 ISRVDDFVGSGIANVILAVAFSPAFARLVRG\ NTLVLKQQTFIESARSIGASDMTVLLRHLIPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMI.NEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRADDFKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNIKTLKLPNFK SCVILLGILLYDVFFVFTFTITKNGESIMVEL AAGPFGNIKKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV OWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVLLVCAGYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRRKEDP-GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	ĺ	1					WGPGAAKSAVGTAILLELVRTFSSMVSNGFR
HIKAVYYSLDNAVLGDDKFHAKTSPLLTSI IESVLKQVDSPNHSQTILYEQVVFTNPSWD\ AEVTRILPMDSSAYSFTAFVGVPAVEFSFME DDQ\AYPFLHTKEDTYENLHKVLQGRLPAVA QAVAQLAGQLLIRISHDRILPLDFGRYGDVV LRRIIGNI.NEFSGDLKARGLTLQWYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIPMGRGDHTLGALL DHI.RLLRSSGTPGATSSTOFQ\ESRFRRQL\ ALL\TWDACKGAANALSGDVWNIDNNF  621 1971 A 4610 793 334 ISRVDDFVGSGIANVILAVAFSPAFARLVRG\ NTLVLKQQTFIESARSIGASDMTVLLRHLIPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMI.NEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRADDFKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNIKTLKLPNFK SCVILLGILLYDVFFVFTFTITKNGESIMVEL AAGPFGNIKKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV VWPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYVSVVTVAYAIGMIL TITVVLGQLMKKGQPALLYLVPCTLITIA/CQFV OWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVLLVCAGYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRRKEDP-GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH				}			PRRSLLFISWDGGDFGSVGSTEWLEGYLSVL
IESVLKQVDSPNHSGQTL YEQVVFTNIPSWD\    AEVIRPLPMDSSAY\SFTAFVGVPAVEPSFME    DDQ\(\text{DDQ\(\text{AYPPLHTKEDTYENLHKVLQGRLPAVA}\)   QAVAQLAGQLLRLSHDRLLPLDFGRYGDVV    LRHIGNLNEFSGDLKARGLT1QWVYSARGDY    IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV    EFYFLSQYVSPADSPFRHIFMGRGDHTLGALL    DHLRLLRSNSSGTPGATSSTGFQESSFFRQL\    ALLIVWDACKGAANALSGDVWNIDNNF    621				i I			HLKAVVYVSLDNAVLGDDKFHAKTSPLLTSL
AEVIRPLPM/DSSAYSFTAFVGVPAVEFSFME	i			l 1			IESVLKQVDSPNHSGOTLYEOVVFTN\PSWD\
DDQ\AYPILHTKEDTYENLHKVLQGRIPAVA QAVAQLAGQLLIRLSHDRLLPLDFGRYGDVV LRHIGRI\neFSGDLKARGLTLQWVYSARGDY IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV EFYFLSQYVSPADSPFRHIEMGRGDHTLGALL DHLRLLRSNSSGTPGATSSTGFQ\ESFFRRQL\ ALL\TWDACKGAANALSGDVWNIDNNF\ 621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRG\ NILVLKQQTFIESARSIGASSMTVLLRHILPGT GSSVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWLQDLIGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFTTPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSV\TVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPFWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLPSKY QYYDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	l			[ [			AEVIRPLPM/DSSAY\SFTAFVGVPAVEFSFMF\
QAVAQLAGQLIRLSHDRILPLDFGRYGDVV		i					DDQ\AYPFLHTKEDTYENLHKVLOGRLPAVA
LRHIGNLNEFSGDLKARGLTLQWYYSARGDY   RAAEKLRQEIYSSEERDERLTRMYNVRIMRV		i i		ì			
IRAAEKLRQEIYSSEERDERLTRMYNVRIMRV							LRHIGNLNEFSGDLKARGLTLOWVYSARGDY
621 1971 A 4610 793 334 ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRGN NTLVLKQQTFIESARSIGASDMTVLLRHILPGT GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDIGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYTYYVSVATVAYAIGMIL TFVVLGUMKKGQPALLYLVPCTLITTA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRIQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQFESANDPEPLCAVCGOAH		1 1		1 1			IRAAEKLROEIYSSEERDERLTRMYNVRIMRV
DHLRILRSNSSGTPGATSSTGFQ\ESRFRRQL\   ALL\TWDACKGAANALSGD\WNIDNNF    621							EFYFLSOYVSPADSPFRHIFMGRGDHTLGALL
ALLITWDACKGAANALSGDVWNIDNNF		1 1		] ]			DHLRLLRSNSSGTPGATSSTGFO\ESRFRROL\
1971   A   4610   793   334   ISRVDDFVGSGIANVIIAVAIFSIPAFARLVRGI   NTLVLKQQTFIESARSIGASDMTVLLRHILPGT   GSSIVVFFTMRIGTSIISAASLSFLGLGAQPFTP   EWGAMLNEARADMVIAPHVAVFPALAIFLTV   LAFNLLGDGLRDALDPKIKG     622   1972   A   4614   2   820   LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ   CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF   RNEDR WAWILQDILGIAFCLNLIKTLKLPNFK   SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL   AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV   VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL   LIAYCRRFDVQTGSSYTYVSVLVAYAIGMIL   TFVVLGILMKKGQPALLYLVPCTLITA/CQFV   AWETVREMKKFWERVTS     623   1973   A   4619   17   691   TLVSVVEFVRADLTREDLAPSSVDSGQAGF   GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT   DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ   IGHFLCLVILVYCAEYINEAAAMNWRLFSKY   QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW   KTLNVMTDLKNAQERRKEKKRRKED*GAA   AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG   LKNRCFI   C624   1974   A   4622   164   668   VSCYTALQSIMNQFESANDPEPLCAVCGOAH							
NTLVLKQQTFIESARSIGASDMTVLLRHILPGI GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLYDVFFVFITFFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVITVAYAIGMIL TFVVLGILMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	621	1971	Α	4610	793	334	
GSSIVVFFTMRIGTSIISAASLSFLGLGAQPPTP EWGAMLNEARADMVIAPHVAVFPALAIFLTV LAFNLLGDGLRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFTIPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVVTVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQFESANDPEPLCAVCGOAH		1 1		1		1	NTI.VI.KOOTEIESARSIGASDMTVI I PUII POT
EWGAMLNEARADMVIAPHVAVFPALAÏFLTV LAFNLLGDGLRDALDPKIKG  622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEFWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		- 10		1	ł		
LAFNLLGDGLRDALDPKIKG		1 1			ł	1	
622 1972 A 4614 2 820 LVYVMIAIFCIASAMSLYNCLAALIHKIPYGQ CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSVYTVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1 1		l			
CTIACRGKNMEVRLIFLSGLCIAVAVVWAVF RNEDRWAWILQDILGIAFCLNLIKTIKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	622:	1972	A	4614	7	820	
RNEDRWAWILQDILGIAFCLNLIKTLKLPNFK SCVILLGLLLLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYVSVYTVAYAIGMIL TFVVLG'LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		*****	*	4014	-	020	
SCVILLGLILLYDVFFVFITPFITKNGESIMVEL AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH							
AAGPFGNNEKNDGNLVEATGQPSAPHEKLPV VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYTYYYSVTVAYAIGMIL TFVVLG\LMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH			[		ſ	ſ	
VIRVPKLIYFSVMSVCLMPVSILGFGDIIVPGL LIAYCRRFDVQTGSSYIYYVSV\TVAYAIGMIL TFVVLGLMKKGQPALLYLVPCTLITA/CQFV AWETVREMKKFWERVTS  623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI  624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH				1	ļ		A CRECIDIEVA DOUBLE AND A CRECIDIEVA DOUBLE AND A CRECIDIEVA DOUBLE AND A CRECIDIEVA DOUBLE AND A CRECIDIA DOU
623 1973 A 4619 17 691 TLVSVVEFVRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFFVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1 1	1		1	ł	
623 1973 A 4619 17 691 TLVSVVEFVRADLTREDLAPSSVDSGQAGF GGCCESGLPNTMPSAFSVSSFFVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1	Ì		ļ		
623 1973 A 4619 17 691 TLVSVVEFVRRADLTREDLAPSSVDSQQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEFWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1 1	ļ		Į.	1	LIAYCRRFDVQTGSSYTYYVSV\TVAYAIGMIL
623 1973 A 4619 17 691 TLVSVVEFVRADLTREDLAPSSVDSQAGF GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		j			-3-		
GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	(22	+1000		10.0		1	
GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH	023	1973	Α	4619	17	691	
DWTEPWLMGLATFHALCVLLTCLSSRSYRLQ IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH					ļ		GGCCESGLPNTMPSAFSVSSFPVSIPAVLTQT
IGHFLCLVILVYCAEYINEAAAMNWRLFSKY QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1		ŀ	I		
QYFDSRGMFISIVFSAPLLVNAMIIVVMWVW KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1 1		j	ļ		IGHFLCLVILVYCAEYINEAAAMNWRLFSKY
KTLNVMTDLKNAQERRKEKKRRKED*GAA AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1	İ		-		
AAWSLRPSRPPSAAPSAAVCVAWASFQLTHG LKNRCFI 624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		1	1	ļ	i		
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH		j Í	ſ	-	ſ		AAWSLRPSRPPSAAPSAAVCVAWASFOTTHO
624 1974 A 4622 164 668 VSCYTALQSIMNQPESANDPEPLCAVCGOAH							LKNRCFI
1 TO THE PER PER PER PER PER PER PER PER PER PE	624	1974	A	4622	164	668	
		'''	<b>^</b>	4022	104	006	VICT TALQSIMINQPESANDPEPLCAVCGQAH
		1 1	1	ļ		J	
TPCGHTYCTLCLTNFLVEKDFCPMDRKPLVL		1	-1				
QHCKKSSILVNKLLNKLLVTCPFREHCTQVL		} I	1			-	QHCKASSILVNKLLNKLLVTCPFREHCTQVL
QRCDLEHHFQTSQAWGTHL*SQLLGRLRQED			- 1	1			
CLSPGVHHCSEV	625	1075	<u>_</u> _+	4625	474	455	
625 1975 A 4625 474 473 CFLSPSPLLPPLLLSSSSSSFSFPLPPPTTLLPSTLP	UZ.3	1 1975	A I	4675	474	Δ73	CEL COCOL I DDI I I COCCODCENT DODOWY I DOWN
PPLLIPSS*LSP	-20	1 -2.2	1	1020	′′′		
		1 /	,	1020		7/3	CELSESELLEFELLSSSSSSSSFSFFLFFFFFILLPSTID

- A	1 050 to		1656	1 20 - 10 - 1		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted and	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
001100		1	***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ľ	l	ſ	residue of		
1		1			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ		1	i	peptide	1	/=possible nucleotide deletion, \=possible
			l	sequence		nucleotide insertion
626	1976	Α	4629	249	3	KLKGNECFCYHCNVCIFLMIKK*GLFLC*IYFI
		l				LFFET*SHSFTRLECSGTISAHCSLQLQGSSNSP
Í		[	1			ASASQVAGIAGTHH
627	1977	A	4635	1	301	FFFFETKPFFAPQAGGQGPSRGSLNPLPTGLK
02,	****	••	1000	_	50.	OFSGLTLSRSGNNGPRPPPRVNFGILRGNGVP
i			1			
1			1		•	PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP
			L			ARAYL
628	1978	Α	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR
			1			TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK
1						NKTAKINCN*RPFLFLVCYLL*AELHIGIFIANF
1		l .				YDCIPNKLNEHLWPKLLQSLIFHVDFCGFLHK
1		I	1			VFYICFTEFLLFLYFL*LFIIKVSCSII*CSTICVF
1		1				SYKSFAVIIFFVDNTRFFSFGF
629	1979	<del>  </del>	4660	18	999	
629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV
}	1					KVKTVCQIPLMKEMLKRFQVAVNLAEDTAH
1				Ì		PKLVFSQEGRYVKNTASASSWPVFSSAWNYF
1			1			AGWRNPQKTAFVERFQHLSCVLGKNVFTSG
l l			1			KHYWEVESRDSLEVAVGVCREDVMGITDRS
1		ļ				KMSPDVGIWAIYWSAAGYWPLIGFPGTPTOO
		ł	i			EPALHRVGVYLDRGTGNVSFYSAVDGVHLH
1		Į.				TFSCSSVSRLRPFFWLSPLASLVIPPVTDRK*G
1						1
1	i	1	i		(	FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG
i		l	ŀ			WSIFWVSLTVPFGICPLCASQEAVPWEVGLA
						NGDGTGNFPRRFWEIFL
630	1980	Α	4669	2	358	FFFFFETESHSVAQAGMQWRNLGSLPAPPPGF
}						TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG
i						FHDVGQDGLDLLTS*STPSASQSAEITGVSHC
!						TRLKKIRFAKGHVEFFFESHVE
631	1981	A	4674	953	614	TPIRGTDDEHEECTVOEYSAGKNTCLRPGAV
1 ***	1.701	1	1071	300		AHTCNPCTLGGRGRWIT*GSGVQDQPGPTWQ
			]			NPVFLERRPRALHSSPGLTTQRILWAQGLWV
	1		1		ĺ	
	1.22					GAGSTGCSRGPRGEGVFREG
632	1982	Α	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP
	ł					*LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP
	1					ASGLAPVPTHWTVSELSRSPVATATFC
633	1983	A	4696	1	1365	RTLGMEGERRASOAPSSGLPAGGANGESPGG
						GAPFPGSSGSSALLOAEVLDLDEDEDDLEVFS
J	J .	J	] ]			KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP
		1				DLKDLFITVDEPESHVTTIETFTTYRITKTSRG
		l				
	1	i	1			EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII
	1	l		]		PPLPEKFIVKGMVERFNDDFIETRRKALHKFL
1		l				NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ
		l	1 :		ľ	GPGLLSRMGQTVRAVASSMRGVKNRPEEFM
1		1				EMNNFIELFSQKINLIDKISQRIYKEEREYFDE
1						MKEYGPIHILWSASEEDLVDTLKDVASCIDRC
		Ì				CKATEKRMSGLSEALLPVVHEYVLYSEMLM
		ļ				
1	1	[	[ ]			GVMKRRDQIQAELDSKVEVLTYKKADTDLL
						PEEIGKLEDKVECANNALKADWERWKQNM
		I	j i			QNDIKLAFTDMAEENIHYYEQCLATWESFLT
	<u> </u>	L				SQTNLHLEEASEDKP
634	1984	A	4708	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ
l		I				WISKAVYKHREMCGLTSTGRKSHGLEKDRM
						FPHAIGGSCRAA*RRKTLQFPCYH
635	1985	A	4709	42	341	YIKQPDAKERRRTVHWKKETESEASEITIPPST
""	1,503	<b>'</b> '	''''		U-74	PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL
						WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT
(26	1006		<u> </u>		0.51	SED SECON
636	1986	A	4721	2	351	EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS
			L			RSTTVTVA+LMQKLNLSMNDAYYTVIMKMSS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	İ	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK
(22		<u> </u>	1000			MYFTTPSNHNAYQVDSVQST
637	1987	Α	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
					ĺ	LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
						DSSASSSRAAGITGVHHHAWLIFFFLVETGFL HAG*AGLELLTSGDPPASASRSAGITGVSHHA
					*	RPRETRFL
638	1988	A	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
	.,00	••			332	TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
						YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
Į į				:		GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
1						LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
					}	WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
639	1989	Α	4743	1040	699	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA
						S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
1 .						AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
	1000		den e			TSNHVLYTQEGLRRGKEG
640	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAQN
						WIRRPCLPSGCLKMNREIGPLQHSLCCPGWS
1						QTPGLKAILLRQPPK*LGLQMESHSCPPAWSA MARSRLTATSASQVQAILLPQPPGTTDSCSPS
1						PDHEQQPLSWVLPPPQKDMNPREQQVALGP
						QAAALPWAVWRNDCFPR
641	1991	A	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
						LQLAASPYFSPSWAECPQPVPAGTHATWCLA
				. [		RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
						FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
					·	QRDTDTGVHTGSGTHTHAHTPPEK
642	1992	A	4798	1	487	GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
			1	•		FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
						TWWFGVKFAAGGLGTFHALLNTAVHVVMY
						SYYGLSALGPAYQKYLWWKKYLTSLQLVQF VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
1 1			8			FLLLFLH
643	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
			.,,,,	-	55.	QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
						MVYFVGENNGDSSHNPVLAATGVGLDVAQS
} }						WEKAIRQALMPVTPQASVCTSPGQGKDHSK
					7	Q*ASVCTSPGQGKDHSKQ
644	1994	Α	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
						LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
				ļ	J	AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
1	1005		4005	450	-10.	SYKDIWGWPCLCGVLHAYIPLLV
645	1995	Α	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV
	ļ	ļ	Į	ļ		AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
			-			PLLAGLVAADAVASLLIVGAVFLCARPRRSP
646	1996	A	4817	47	1033	AQEDGKVYINMPGRG LOGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
""	1,,0	^	701/	71	ררויז	LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
] ]		J	j			HKQAVQCLKGPGQVARLVLERRVPRSTOOC
1 1			ĺ			PSANDSMGDERTAVSLVTALPGRPSSCVSVT
			ļ			DGPKF*SSN*KRIANGLGPSFVQMEKESCSHL
		- 1	l		j	KSDLVRIKRLFPGHPAEENGALAAGDIILGRE
'	}	]	ļ		*	WEGPRKASSSRCRGSWAMQLSVQAGPSFAS
[ [		ĺ	- 1			YYPAAVEVLHILLRGAPQEVTLLLCRPPPGAL
			1			PELEQEWQTPELSADKEFTRATCTDSCTSPIL
	1		ļ			GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
647	1997	<del>,  </del>	4854	1044	225	EGTMGAKTERDLGPVP
<u> </u>	177/	A ]	+0.24	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide seq-	seq- uence		USSN 09/496	location correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
nence	dance		914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine,
1 444	ľ	i	711	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł				peptide		/=possible nucleotide deletion, \-possible
1.		ŀ		sequence		nucleotide insertion
						IVMPTYDLTDSVLETMGRVSLDMMSVQANT
		ļ				GPPWESKNSTAVWRGRDSRKERLELVKLSRK
		]	1		1	HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
						FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
			·			QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK
·		i			1	LKWAKDHDEEAKKIAKAGQEFARNNLMGD
648	1998	A	4867	2020	027	DIFCYYFQTFPRNMPIYK
046	1998	Α .	480/	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTQ SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI
		i				
		ì				LTEQHSKRVAVILNEFGEGSALEKSLAVSQG GELYEEWLELRNGCLCCSVKDNGLRAIENLM
1		ł	1			QKKGKFDYILLETTGLADPGAVASMFWVDA
-						ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI
1 .						NEATROVALADAILINKTDLVPEEDVKKLRT
1				,	1	TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL
1			1		1	SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG
· .						NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV
					•	IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
						SWKDDTERTNRLVLLGRNLDKDILKQLFIAT
	1000		10.50			VTETEKQWTTHFKEDQVCT
649	1999	A.	4873	226	189	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR
'						FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
						GQAGLELRTSGDPPASASQSAGITGVSHLA*P TSMPLLPFQRLCVYI
650	2000	A	4874	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF
1 050	2000	^	1074	2	437	K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG
						FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR
						CPASFYLFLKYYLEAKFCA*GECAPSAGVGA
						GYKRGHKSCLLINCVVQI
651	2001	A	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR
						PHMEPKASCPAAAPLMERKFHVLVGVTGSV
						AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
1 '			í I			SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL
						RRWADLLLVAPLDANTLGKVASGICDNLLTC
1 1			ł			VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
						VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA
1 1						EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP LYSEWVQAKSVKMDVGKIGGYPHLLNGGPA
		00				LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA
						SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
	i					PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR
				·		GQSPIPSRASSPSCSWAQVPGVALARCAGVC
	ſ	-	[	ĺ	ĺ	KPGDSWRVAACISGRCCSRGRRRGSGPRNPE
		l			l	QSFRGAWGPSFWGSWKSQRELSAGGAQAWP
L	2002					LLGSAGSGLRGEA
653	2003	Α -	4965	2	283	FFFFI*DGVSLCHPGWNAVARSWLTATSASR
						VQAVSCFRLPSSWDYRHATMPG*FF*YF**R
654	2004	Ā	4968	3	437	WGFTILAILVLNS*PQVICPPWPPKVLTLQA
400	2004	Λ	4700	٠ .	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD
			1		)	IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW DPERDPGCRVQVSRGPRFGEQKTPGLOGCLP
1			ĺ	į		PPCLTHLAAASCVVVWCGRWKRDSAECQCD
[ [	i			ĺ		HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNTTCIQPSMISSMALPITYILLCIVGVFGN
						TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI
/						REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES
			<u> </u>			

NO: of   NO: of   No: of   note   n	CEOTO	SEQID	Met	SEQ	Predicted	Deadist dead	Amino acid sequence (A=Alanine C=Cysteine,
Description   Description	SEQ ID					Predicted end	
			поа				
Sequence		1	1				
Beautiful			1				
amino acid residue of peptide sequence		uence	1				
Peptide   Pept	uence			914			
Pepilide   Sequence				1			T=Threonine, V=Valine, W=Tryptophan,
						sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		1	1.	1 * *		
AIWWECKROWLLQTHWTLDKYGLADARI   GFGP(RRPVLR,PRRARLIX,)					sequence		
FGPOHRPVILRLPMRALLX*   558   2008   A   5017   1   292   PFPFKETESHSYTQAGVQWINDGSLOPPPE   KRRSCLSLLSSWDYRCAPPHP ANFVPLVETC   HHVAQAGLKLTT.*SANLGLSTSLPDFFILL   HHVAQAGLKLTT.*SANLGLSTSLPDFFILL   HHVAQAGLKLTT.*SANLGLSTSLPDFFILL   HHVAQAGLKLTT.*SANLGLSTSLPDFFILL   HHVAQAGLKLTT.*SANLGLSTSLPDFFILL   HTVSLVSPVLGKWSPCLQGGLSAVHTVPN   MAACWAVHYKTHMPROLAVLPRLVLNSW   MAACWAVHYKTHMPROLAVLPRLVLNSW   *AIIILWPFKALGLQA   HFGHLCLQPFGGPA*FLDYCRGCCPIPYCRGC   HFGHLCLQPFGGPA*FLDYCRGCCPIPYCRGEP   HHTGARWNH   ROSCFSTGAKYQWFHYOPLQSCPPGILKQSS   AGSCPRQKKTTPOFTVLCVCSFWTYQRGEP   HHTGARWNH   ROSCFSTGAKYQWFHYOPLQSCPPGILKQSS   LSLPNSRDHRHVPPRLAIFSPAETGSPYFAQA   LELLGSSHPTSASQSARITGVSHRAWPLK**   NINQYQTLTMR   NINQYQTLTMR   NINQYQTLTMR   NINGYQTTA   NINGYQTTA   NIN			1			1	VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
658   2008   A   5017   1   292   FPFFKETESISYTOAGQWEIDLGLEPPER   FROM   FPFFKETESISYTOAGGWEIDLGSLEPPER   FPFKETESISYTOAGGWEIDLGSLEPPER   FPFKETESISYTOAGGWEIDLGSTSLEPPER   FPFKETESISYTOAGGCARPIPANFYTI-VETCO   FPFKETESISYTOAGGCARPIPANFYTI-VETCO   FPFKETESISYTOAGGCARPIPANFYTI-VETCO   FPFKETESISYTOAGGCARPIPANFYTI-VETCO   FPFKETESISYTOAGGCARPIPANFYTI-VETCOAGGCARPIPANFAGARA   FPFKETESISYTOAGGCARPIPANFAGARA   FPFKETESISYTOAGGCARPIPANFAGARA   FPFKETESISYTOAGGCARPIPANFAGAGARA   FPFKETESISYTOAGGCARPIPANFAGAGAGARA   FPFKETESISYTOAGGAGARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISYTOAGGAGAARA   FPFKETESISTOAGAARA   FPFKETESIAGGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPFKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAARA   FPKETESISTOAGAAARA   FPKETESISTOAGAAAAAA   FPKETESISTOAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		l	1	1			AIWWEQKRQWLLQTHWTLDKYGILADARLF
KRPSCL SLLSSWDYRCAPPIPANFVTL VETCH			ľ	1	1		FGPQHRPVILRLPNRRALRLX*
KRPSCLSLLSSWDYRCAPPIPANFUTUETC	658	2008	Α	5017	1	292	
HHYAQAGLKLITI-\$ANIGLSTSLPPLFILL   659   2009   A   5018   17   338   RGHGKSLTGGTFGRWGDGLLVSEDWSHL   TYNSLVSPVLGKWSPCLQFGLSAVHTWPH   MAACWAVHVKTHMRPGLAVLPRLVINSW   *AIIILWPPKAGLQQA   *AIIILWPFKAGLQQA   *AIIILWPFKAGLQQA   *AIIILWPFKAGLQQA   *AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPF   HRTQARWNH   *AGSCSPRQKKTTPGPTVLCVCSFWIYQRGEPF   HRTQARWNH   *GSCSSTQAKVQWFHYGPLQSQPPGLKQSS   LSLPNSRDHRRIVPPRLAIFSFAETGSPYFAQA   LELIGSSIPPTSASQSARITGVSHRAWPLK*I   NLNQYQTLTMN   *ASSCREPTSASQSARITGVSHRAWPLK*I   NLNQYQTLTMN   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSARITGVSHRAWPLK*I   *ASSCREPTSASQSTASALLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSQSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNAAITIDG   *TENNACSGSPSCAELLEAGANNA							
	l	]					
TYNSLVSPVLOGKVSPCLOGPGLSAVHTWFW   MAACWAVHYKTHMRPGIAVLPRLVINSW   MAACWAVHYKTHMRPGIAVLPRLVINSW   AIILL WPPKALGLQA	650	2000	<u> </u>	5018	17	229	
MAACWAYHYKTHARPGLAVLPRLVLNSW   AIILLWPPKALGLQA	037	2009	^	3016	17	330	
*AIIILWPPKALGI.QA		1	1	1	1		
660   2010   A   5028   2   310   SRVDDFVGERRGCDELICGHRGLRAUPLE   HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGG   AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPHRTGARWH   RQSCSSTQAKVQWFHYGPLQSQPPGLKQSS   LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQA   LELLGSSHPPTSASQSARITOVSHRAWPLK*I   NLNQYQTLTMN   RQSCSSTQAKVQWFHYGPLQSQPPGLKQSS   LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQA   LELLGSSHPPTSASQSARITOVSHRAWPLK*I   NLNQYQTLTMN   EANSQRILLAURILLSQGYNVNAVTLDHVT   LHEACLGDHVACARTILLEAGANVNAITIDG   TPLFNACSQGSPSCAELLLEYGAQAQLESCL   SPTHEGASKGHHECLDILISWGIDVDQEPPHS   TPLYVACMAQQFHCIWNLIYAGAGVSKGK   WDTPLPGAGHGSTQLE-IFAMVEIWQ   WDTPLPGAGHGSTQLE-IFAMVEIWQ   WDTPLPGAGHGSTQLE-IFAMVEIWQ   WDTPLPGAGHGSTQLE-IFAMVEIWQ   WDTPLPGAGHGSTQLE-IFAMVEIWQ   WDTPLPGAGHGSTAL-INAAFPHLVHAELHPSS   WLRGLLGVVGAAVAAVLSFLYRVLVKSQ   GPAEPSTITIFQLATMAAFPHLVHAELHPSS   WLRGLLGVVGAAVAAVLSFLYRVLVKSQ   GPAEPSTITIFQLATMAAFPHLVHAELHPSS   WLRGLLGVVGAAVAAVLSFLYRVLVKSQ   GPAEPSTITIFQLATMAAFPHLVHAELHPSS   WLRGLLGVVGAAVAAVLSFLYRVLVKSQ   TSFLGCAIQMFSFLYFUKMVDLLSQKK   TSFLGCAIQMFSFLYFUKMVDLLSQKK   TSFLGCAIQMFSFLYFGSSLSRL   QLLFVIFLLYLFFISSNGHEDSVLDLSQKK   TSFLGCAIQMFSFLYFGSSHSFLLAAMGYDR   YMAICNPLRYSVLMGHGVCMGIMAAAWA   GFVSLVTTSLVFULPFHSSNGHEDS   DLVQRSLYCDTAVNDVLNTSVTSMGSQMP   HDQNEGFHCREECRILGHSDRCWMPRNPMP   RSKSPEHVRNILIAISETAAADVEAVDCOPP   KRTFATFGKDVSDHPAEERPTLKGKRTVDV   ICSPKVNSVIREAGNGCCAISPSTSPLSHLKSSI   PTLYSVSYEIDDPGTTARC   MILLSTSS*VYFQSSTRDSHFTLFDFQKTGFP   CFSKVNSYEIDPGTTARC   MILLSTSS*VYFQSSTRDSHFTLFDFQKTGFP   CFSKVNSYEIDPGTTARC   MILLSTSS*VYFQSSTRDSHFTLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSYEIDPGTTARC   DLTNSHFFLFDFQKTGFP   CFSKVNSTEMKRLNRLLDNS   TCADQSVIWKLSEDKQLAICLKYAGVHAEN   EDVEGROVFNTKPIAQLIELAKSNPQQVV   CCSDMATTFNGLTPGMEVMMYGLYPLRAU   EDVEGROVFNTKPIAQLIELAKSNPQQVV   CCSDMATTFNGLTPGMEVMMYGLYPLRAU   EDVEGROVFNTKPIAQLIELAKSNPQQVV   CCSDMATTFNGLTPGMEVMMYGLYPLRAU   CCSDMATTFNGLTPGMEVMMYGLYPLRAU   CCSDMATTFNGL	J	1	)	1	}	}	
HPGHLCLQPPGGPA*FLDYCRGCCPHPVPGS AGSCPRQKKTTPGFTVLCVCSFWTYQRGEPI HTGARWNH		l	ļ.,				
AGSCPRQKKTTFGPTVLCVCSFWIYQRGEP:   HRTGARWNH	660	2010	A	5028	2	310	
HRTGARWNH			1				
661   2011   A   5050   752   431   RQSCSSTQAKVQWFHYGPLQSQPPGLKQSS LSLPNSRDHRHVPPRLALIFSAETGSPYFAQA LELLGSSHPPTSAQSARITGVSHRAWPLK*1   NLNQYQTLTMN	l	1	1	1	ł	}	AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH
LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQA   LELLGSSHPPTSASQSARITGVSHRAWPLK*    NINDYQTLTM    NINDYQTLTM    NINDYQTLTM		1		'			HRTGARWNH
LSLPNSRDİRHYÖPRLAIFSÄĞETĞSPYÄQA   LELLGSSHPTTSASQARITGVSHRAWPLK*I   NLNQYQTLTMN	661	2011	Α	5050	752	431	ROSCSSTOAKVOWFHYGPLOSOPPGLKOSSO
LELIGSSHPTSASQSARITGVSHRAWPLK*							
NILNOYQILTIMN	ł	1	1		•		
662   2012   A   5054   48   103		J	}	1			NI.NOYOTI.TMN
EAASQGRLALRTLLSQGYNVNAVTLDHVT	662	2012	A	5054	48	103	
LHEACLGDHVACARTILEAGANVNAITIDG: TPLFNACSQGSPSCAELLLEYGAQAQLESCL SPTHEGASKGHHECLDILISWGIDVDQEPHS TPLYVACMAQQFHCIWNLIYAGAGVRKGK: WDTPLPGAGHQSTQKLE*LFAMVEIWQ WDTPLPGAGHQSTQKLE*LFAMVEIWQ VRNS*SPAHCASVYKHHYMDGQTPCLFVSS ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSC. GPAEPSTTIFTQLATMAAFPHLVHAELHPSSS WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ LSFIEVLSMEQVNKTVVREFVVLGFSSLARL-QLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTI MYFFLAILSCSEICYTFVIVFKMLVDLLSQKK TISFLGCAGMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAGFTVSLVTTSLVFHLPFHSSNQHE QQYHNTGSGGHHAHCQVGHSPHVHYPSGC PL*QRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPH HDQNEGFHCREECRILGHSDRCWMPRNPMF RSKSPEHVRNIIALSIEATAADVEAYDDCGFT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGFPT VGFKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGFTGFT VGFKAQLSGLQLQPCLYKRR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQFFYLGHT IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVTWKLSEDKQLAICLKYAGVHAEN. EDYTEGGDLYVTVTEGGIVLSRELMKRLNRLLDNST TCADQSVTWKLSEDKQLAICLKYAGVHAEN. EDYTEGGDLYSVTVEGGIVLSRELMKRLNRLLDNST TCADQSVTWKLSEDKQLAICLKYAGVHAEN. EDYTEGGTUSRELMKRLNRLLDNST TCADQSVTWKLSEDKQLAICLKYAGVHAEN.	002	2012	1	3034	70	'03	
TPLFNACSQGSPSCAELLLEYGAQAQLESCL   SPTHEGASKGHHECLDILISWGIDVOGEPHS   TPLYVACMAQQFHCIWNLIYAGAGYRKGKY   WDTPLPGAGHQSTQKLe*LFAMVEIWQ   WDTPLPGAGHQSTQKLe*LFAMVEIWQ   WDTPLPGAGHQSTQKLe*LFAMVEIWQ   WDTPLPGAGHQSTQKLe*LFAMVEIWQ   GPAFPSTITFTQLATMAAFPHLVHAELHPSSS   WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ   GPAFPSTITFTQLATMAAFPHLVHAELHPSSS   WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ   ULSFEVLSMEQVNKTVVREFVVLGFSSLARL   QLLFVIFLLLYLFTLGTNAIIISTTVLDRALHTI   MYFFLALISCSEICYTFVIVPKMLVDLLSQKK   TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR   YMAICNPLRYSVLMGHGVCMGLMAAAWA   GFTVSLVTTSLVFHLPFHSSNOHE   G65   QQYHNTGSAGHHAHCQVGHSPHVHYPSGC   FL*IQRGLPS*NSLEGHSLKDSGHEESVQLDS   HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI   HDQNEGFHCREECRILGHBDRCWMPRNPMP   RSKSPEHVRNIILALSEATAADVEAYDDCGPI   KRTFATFGKDVSDHPAEERPTLKGKRTVDV   ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI   PTKPSVSYEIVDPGITARC   G66   2016   A 5080   408   248   IMLLSTSS*VYPQSSTKDSHFFLFDFQKTGPPI   VGPKAQLSGLQLQPCLYKRR   G67   2017   A 5081   129   247   DLTNSHFTLFDFQKTGPPLGGPKAQFSSLQLC   PCVY*RR   G68   2018   A 5086   852   233   NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF   RALPTITADIENLKYLLFTRDASQPFYLGHT   IFGDLEYVTVEGGIVLSRELMKRLNRLLDNS   TCADQSVIWKLSEDKQLAICLKYAGVHAEN   EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEGCSDMAITFNGLTPQKMEVMMYGLYRLRAK   EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEGCSDMAITFNGLTPQKMEVMMYGLYRLRAK   EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEGCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPQKMEVMMYGLYRLRAK   ECCSDMAITFNGLTPGLAKYLLFTLANGT   ECCSDMAITFNGLTPGLYRAC   ECCSDMAITFNGLTPGLYRAC   ECCSDMAITFNGLTPGLYRAC   ECCSDMAITFNGLTPGLYRAC   ECCSDMAITFNGLTPGLYRAC							
SPITHEGASKGHHECLDILISWGIDVDQEIPHS		ļ	1	1	ļ	!	
TPLYVACMAQQFHCIWNLIYAGAGVRKGKT   WDTPLPGAGHQSTQKLE*1FAMVEIWQ							
WDTPLPGAGHQSTQKLE*LFAMVEIWQ		ļ.	ļ				
663   2013   A   5066   951   580   VRNS*SFAHCAŠVYKIHYMDGQTPCLFVSS ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSC GPAEPSTIFTQLATMAAFPHLVHAELHPSSI WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ GPAEPSTIFTQLATMAAFPHLVHAELHPSSI WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ (LJFVIFLLLYLFTLGTNAIIISTIVLDRALHTI MYFFLAILSCSEICYTFVVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAGFTVSLVTTSLVFFILPFHSSNQHE (PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC (MLLSTS*VYFQSSTKDSHFFLFDFQKTGPPI VGFKAQLSGLQLQPCLYKRR (MLSTS*VYFQSSTKDSHFFLFDFQKTGPPI VGFKAQLSGLQLQPCLYKRR (MLSTS*VYFQSSTKDSHFFLFDFQKTGPPI VGFKAQLSGLQLQPCLYKRR (MLSTAYKRH) DLTNSHFFLFDFQKTGPPI GPKAQLSGLQLQPCLYKRR (MLSTAYKRH) RALPTITADIENLKYLLFTRDASQFYLGHT IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRRAE		į	1				
ADLPEGVAVSGPSPAEFCRKHRLPAPVPPSC GPAEPSTTIFTQLATMAAFPHLVHAELHPSSI WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ ULFVIFILLYLFTLGTNAIIISTIVLDRALHTI MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWA GFTVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMS HDVQRSLYCDTAVNDVLNTSVTSMGSQMS HDVQRSLYCDTAVNDVLNTSVTSMGSQMS RSKSPEHVRNILALSIEATAADVEAYDDCGFT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMILSTSS*VYPQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLG PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRRAE			<u> </u>	<b></b>			
GPAEPSTTIFTQLATMAAFPHLVHAELHPSSE WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ  664 2014 A 5071 550 1 LSFIEVLSMEQVNKTVVREFVVLGFSSLARL QLLFVIFLLLYLFTLGTNAIISTIVLDRALHTI MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAI GFITVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGFT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMILSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFIRDASQFFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVER CCSDMAITFNGLTPQKMEVMMYGLYRLRAI	663	2013	A	5066	951	580	
WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ							
1			1				GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
1			İ	i			WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
QLLFVIFLLIYLFTLGTNAIIISTIVLDRALHTI MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAG GFTVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNBGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPI KRTFATFGKDVSDHFAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPIGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFIRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLO
MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAG GFTVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTTATFGKDVSDHPAEERPTLKGKRTVDV' ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFIRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEC CCSDMAITFNGLTPQKMEVMMYGLYRLRAF			1				
TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAG GFTVSLVTTSLVFHLPFHSSNQHE  665  2015  A 5074  496  692  QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC  666  2016  A 5080  408  248  IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667  2017  A 5081  129  247  DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668  2018  A 5086  852  233  NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF							
YMAICNPLRYSVLMGHGVCMGLMAAAWAGGFTVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCGPL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSHDVQRSLYCDTAVNDVLNTSVTSMGSQMPIHDVQRSLYCDTAVNDVLNTSVTSMGSQMPIHDQNEGFHCREECRILGHSDRCWMPRNPMPRSKSPEHVRNILALSIEATAADVEAYDDCGPTKRTATFGKDVSDHPAEERPTLKGKRTVDVICSPKVNSVIREAGNGCEAISPVTSPLHLKSSLPTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPIVGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQPCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVFRALPTTADIENLKYLLFTRDASQPFYLGHTVIFGDLEYVTVEGGIVLSRELMKRLNRLLDNSTTCADQSVIWKLSEDKQLAICLKYAGVHAENLEDYSGDFTNTKPLAQLIEEALSNNPQQVVECCSDMAITFNGLTPQKMEVMMYGLYRLRAF		1	1 ·	ſ			
GFTVSLVTTSLVFHLPFHSSNQHE  665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF				1			
665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCO PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF		ļ					
PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDS HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMILSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	665	2015	-	5074	406	602	
HDVQRSLYCDTAVNDVLNTSVTSMGSQMPI HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERFTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEC CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	003	2013	^	3074	470	092	
HDQNEGFHCREECRILGHSDRCWMPRNPMP RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSI PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQFFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEC CCSDMAITFNGLTPQKMEVMMYGLYRLRAF						i	
RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMILSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPLAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF			1				
KRTFATFGKDVSDHPAEERPTLKGKRTVDV ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPLAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF		1					
ICSPKVNSVIREAGNGCEAISPVTSPLHILKSSL PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTITADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF		1					
PTKPSVSYEIVDPGITARRC  666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	·	[	ĺ				
666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPI VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTADSQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPLAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF				] ,			ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNST TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF		<u></u>		<u> </u>			
VGPKAQLSGLQLQPCLYKRR  667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNSY TCADQSVIWKLSEDKQLAICLKYAGVHAENL EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	666	2016	Α	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLC PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRSLDNS) TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF							
PCVY*RR  668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN, EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	667	2017	Α	5081	129	247	DLTNSHFFLFDFOKTGPPLGGPKAOFSSLOLO
668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN, EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF				-	-		
RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSI TCADQSVIWKLSEDKQLAICLKYAGVHAEN EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	668	2018	A	5086	852	233	
IFGDLEYVTVEGGIVLSRELMKRLNRLLDNS: TCADQSVIWKLSEDKQLAICLKYAGVHAEN. EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEC CCSDMAITFNGLTPQKMEVMMYGLYRLRAF	300		١ ' `	""			
TCADQSVIWKLSEDKQLAICLKYAGVHAEN, EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEC CCSDMAITFNGLTPQKMEVMMYGLYRLRAF					l	ļ	· · · · · · · · · · · · · · · · · · ·
EDYEGRDVFNTKPIAQLIEEALSNNPQQVVE CCSDMAITFNGLTPQKMEVMMYGLYRLRAF			1	]	l		
CCSDMAITFNGLTPQKMEVMMYGLYRLRAF			l	1 1	ļ	į	
					l		
	1				l		
GHYFNDTLVFLPPVGSEND							GHYFNDTLVFLPPVGSEND
669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG	669	2019	A	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPI							ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
RGCQHEAAPCPRGPGSDGLHHASAACASLP7							RGCQHEAAPCPRGPGSDGLHHASAACASLPP
							SPILPVLLPELGPL
			A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
						C 40	TO A TOTAL OF THE PARTY OF THE

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DEL*NMNGRVDYLVTEEEINLTRGPSGLGFNI VGGTDQQYVSNDSGIYVSRIKENGAAALDGR LQEGDKILSVNGQDLKNLLHQDAVDLFRNA GYAVSLRVQHRLQVQNGPIGHRGEGDPSGIPI
671	2021	A	5105	672	400	FMVLVPVFALTMVAAWAFMRYRQQL RDGREELCLQQEPILPSRICSSAPLLYFLFICPF VLLLLLLISLLCLYWKARKLSTLRSNTRKEKA LWVDLKEAGGVTTNRMED*EEDECN
672	2022	A	5148	72	314	IIYFSYNIFLKITELLNDVERLKQALNGLSQLT YTSGNPTKRQSQLIDTLQHQVKSLEQQLAVS NQAHGALQEYVLAPCS
673	2023	A	5152	210	335	REILCSRIGRLNIV*MSLFPNLTCRLNAIPIKIPA NHFVEVT
674	2024	A	5153	3	2953	LTEDOPFDILQKSLQEANTTEQTLAEEAYLDA SIGSSQOFAQAQLHPSSSASFTQASNVSNYSG QTLQPIGVTHVPVGASFASNTVGVQHGFMQH VGISVPSQHLSNSSQISGSGQIQLIGSFGNHPS MMTINNLDGSQIILKGSGQQAPSNVSGGLLV HRQTPNGNSLFGNSSSSPVAQPVTVPFNSTNF QTSLPVHNIIQRGLAPNSNKVPINIQPKPIQM GQQNTYNVNNLGIQQHHVQQGISFASASSPQ GSVVGPHMSVNIVNQQNTRKPVTSQAVSSTG GSIVIHSPMGQPHAPQSQFLIPTSLSVSSNSVH HVQTINGQLLQTQPSQLISGQVASEHVMLNR NSSNMLRTNQPYTGPMLNNQNTAVHLVSGQ TFAASGSPVIANHASPQLVGGQMPLQQASPT VLHLSPGQSSVSQGRPGFATMPSVTSMSGPSR FPAVSSASTAHPSLGSAVQSGSSGSNFTGDQL TQPNRTPVPVSVSHRLPVSSSKSTSTFSNTPGT GTQQQFFCQAQKKCLNQTSPISAPKTTDGLR QAQIPGLLSTTLPGQDSGSKVISASLGTAQPQ QEKVVGSSPGHPAVQVESHSGGQKRPAAKQ LTKGAFILQQLQRDQAHTVTPDKSHFRSLSD AVQRLLSYHVCQGSMPTEEDLRKVDNEFETV ATQLLKRTQAMLNKYRCLLLEDAMRINPPAE MVMIDRMFNQEERASLSRDKRLALVDPEGFQ ADFCCSFKLDKAAHETQFGRSDQHGSKASSS LQPPAKAQGRDRAKTGVTEPMNHDQFHLVP NHIVVSAEGNISKKTECLGRALKFDKVGLVQ YQSTSEEKASRREPLKASQCSPGPEGHRKTSS RSDHGTESKLSSILADSHLEMTCNNSFQDKSL RNSPKNEVLHTDIMKGSGEPQPDLUTKSLET TFKNILELKKAGRQPQSDPTVSGSVELDFPNF SPMASQENCLEKFIPDHSEGVVETDSILEAAV NSILEC
675	2025	A .	5154	599	1880	LKKMEPFSCDTFVALPPATVDNRIIFGKNSDR LYDEVQEVVYFPAVVHDNLGERLKCTYIEID QVPETYAVVLSRPAWLWGAEMGANEHGVCI GNEAVWGREEVCDEEALLGMDLVRLGLERA DTAEKALNVIVDLLEKYGQGGNCTEGRMVF SYHNSFLIADRNEAWILETAGKYWAAEKVQE GVRNISNQLSITTKIAREHPDMRNYAKRKGW WDGKKEFDFAAAYSYLDTAKMMTSSGRYCE GYKLLNKHKGNITFETMMEILRDKPSGINME GEFLTTASMVFILPQDSSLPCIHFFTGTPDPER SVFKPFIFVPHISQLLDTSSPTFELEDLVKKKS HFKPDRRHPLYQKHQQALEVVNNNEEKAKI MLDNMRKLEKELFREMESILQNKHLDVEKIV NLFPQCTKDEIQIYQSNLSVKVSS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide seq-	peptide seq- uence		in USSN 09/496	nucleotide location correspondi	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
испсе			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
676	2026	A	5155	sequence	306	nucleotide insertion  FFFLRRSLALSPRPDCGLQWRNLGSLQAPPPG
0.0	2020		3133		300	FTPFSCLSLPSSWDYRRPPPRPANFLYF**RRG FTLLARMVSIS*PHDPPASASQSAGITGVSHRA RPT
677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR
						LNKRSFFMISPTDQQVHCWAWLKKHMPKDS NLLLEDVTWKYTALNLIGPRAVDVLSELSYA PMTPDHFPSLFCKEMSVGYANGIRVMSMTHT
678	2028	A	5183	1919	2018	GEPGFMLYIPIEYRWGFTMLSTLVSNS PALCRLRDDMTVCVADFGLSKKIYSGDYYRO
		••	7103	.,,	2010	GRIAKMPVKWIAIESLADRVYTSKSDVWAFG VTMWEIATRGMTPYPGVQNHEMYDYLLHG
679	2029	Α	5190	39	499	HRLKQPEDCLDELCKI**SPQSP RESOVKHFKMRKIDLCLSSEGSEVILATSSDE
						KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH VRIERLVIQSYFVQTLKIEKSTSKEPVDFEQWI
						EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS
680	2030	A	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG
681	2031	A	5207	10	247	LLSRKASHCYCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI
						KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK
682	2032	A	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS
683	2033	A	5218	85	402	CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPTDQEPKTVMLSK QSESAI
684	2034	A	5220	1	194	NLMKEMONLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPOHL
685	2035	A	5228	260	440	TDS  LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED  QKKTPQVPINFTELSKCS*S*KIMSGERE
686	2036	A	5239	79	508	GGEAAARAKLSSPRPHRVGRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD
687	2037	A	5244	1	428	ESCKKHLGRVLSIWEERS MAAVVAATALKGRGARNARVLRGILAGATA
						NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR
688	2038	A	5249	1	1407	ASSQSAPSPDVGSGVQT LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS
			,			EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS
						KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

NO. of nucle ordinal control of the property o	SEQ ID	I SEO ID	Met	SEQ	Predicted	1 20-27-4-4	
nucheotide sequence corresponding plate unione corresponding to the state of the sequence period in the sequence corresponding a first first corresponding plate and sequence period for peptide sequence period for period for peptide sequence period for pe		1 -				Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
comide sequence unice unice unice sequence seque	1		1100				D=Aspartic Acid, E=Glutamic Acid,
seq. uence		1	1				r=Phenylalanine, G=Glycine, H=Histidine,
uence    914   anito first anito of peptide periodic of peptide sequence   Peptide sequen			1				i=isoleticine, K=Lysine, L=Leucine,
amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per seq		acite	1				M=Methionine, N=Asparagine, P=Proline,
residue of popular sequence   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown, *=Sup codon,   Y=Tyrosine, X=Unknown,   Y=Tyros	uence	ł	1	914			Q=Glutamine, R=Arginine, S=Serine,
Peposible nucleotide cletion, 1-possible nucleotide insertion   Peposible nucleotide insertion   PSSTMSLAKLLORERGISARYVISPISEMPLOPI, PSSTMSLAKLLORERGISARYVISPISEMPLOPI, PSSTMSLAKLLORERGISARYVISPISEMPLOPI, PSSLARSTPHNSPENSCYNLERPSYPULSEN   FLASRPAFITH, OBMYGLRBRAMPPOVGOLAM   RIVERISARY AND PASSA   PSSTMSLAKULORERGISARYVISPISMSCHOOLER   PSSTMSLARYVINSGASPARYPOVGOLAM   RIVERDAM   RIV		J	J	1			
	1			1	1	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
FISTIMILAKLIQERGISAKVYHSPISENPILOPH PKSLAJRYSPINFSHEPSPCPSH PERSPEVHLSSN FLASRPASTRI QBMYGLRSRNPPDVQULKM NLVDRLKRLGIARVVKNPGAQBERGOEAEI GPQKPDSAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNSGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNGSSLLGGLRRNQSLPVIM GSPAAPVCTSSYMAVLKED RYNAVALNGSSLLGGLRRNGSLPVIM RYNAVALNSGSSLLGGLRRNGSLPVIM RYNAVALNGSPERVSYNGKFGEVINSMMSPEKKMYKKFGS GAPGAGAGGAKANSPPEVKVNRQSFQLIG RKTRIBVGLAFVSNMSPEEKMYKKFGSLLAKSK GEKRERQAQREDALELTEKLDQDWKEIQTILL SHKTYKSENDOKKEKPFDAYDMWNELGF EMKAQPSNRMSTEALLAKEGGEHLRKLBAE RRRNGKRENGSSEGDETESDSPPSHLDLES NVESEENEKPAKEQROTPGKGLISGKERAG KATRDELPYTFAAPSYEELSSLLGRSMEQ LLVVERIQKCNIPSLAGGNKAKLEKLFGFLLE SPAESPEEGDSSGGEDTESDSPPSHLDLES NVESEENEKPAKEQROTPGKGLISGKERAG KATRDELPYTFAAPSYEELSSLLGRSMEQ LLVVERIQKCNIPSLAGGNKAKLEKLFGFLLE SPAESPERPYTAMPTALVCL GQLTKCPILSLQDVVKGLFVCCLFLEVYALS QRIPELINFLIGH LYMFNAKSQGSTLVHPR ALGKNSELLVSAREDVATWQOSSLSLRWA SSLRAPTSTEANHRISCLAVGLALLHCOMP ESASDAIKFVLRDAMHEMEEMETKGRAALP GSADAIKFVLRDAMHEMEEMETKGRAALP GSADAIKFVLRDAMHEMEEMETKGRAALP GSADAIKFVLRDAMHEMEEMETKGRAALP GRIPELINFLIGH LYMFNAKSGSSLUVHPR ALGKNSELLVSAREDVATWQOSSLSLRWA SSLRAPTSTEANHRISCLAVGLALLHCOMP ESASDAIKFVLRDAMHEMEEMETKGRAALP GRIPELINFLIGH LYMFNAKSGSSLUVHPR ALGKNSELLVSAREDVATWQOSSLSLRWA SSLRAPTSTEANHRISCLAVGLALLHCOMP ESASDAIKFVLRDAMHEMEEMETKGRALP VSQLAFTSTEANHRISCLAVGLALLHCOMP ESASDAIKFVLRDAMPALLTDHLADCSHOPLU GRIPTINFLICHTURGAGGELTAFTQESTDPTW ASQSAGITOVSHCSWPVIYVLSTLIHAVRNVL FRITEPISKSSFLSYDKEIPILVVLKFYLVTLT FRYTVENSSELSTYPLISTLEAVCVCRLWRGCVRVLR FRYTVLYSTAGGAGGETTSLECKGGNOTTGRIFTLRCN CHOPP CONTRACTOR OF THE TOTAL	1	1	l	ł		ł	
PRSLAPSTPPNSPHSPCRSPLPERVHUSEN   FLASRPAETTIC, QUYGLERS   FLASRP AGENGROCA     FLASRP AGENGRAVION   FLASR AGENGROCA     GPOKEPSAVILING SSSLIGGLERNOS, LEVIM     GSFAAP VCTSSTRMGVLKED     SEAR AGENGRAMAKAKK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GALHKAT VGALHE KAMAKA     GALHKAT VGALHA GARGAK AL KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GA	<b>}</b>	<u> </u>	<del> </del>		sequence		
PRSLAPSTPPNSPHSPCRSPLPERVHUSEN   FLASRPAETTIC, QUYGLERS   FLASRP AGENGROCA     FLASRP AGENGRAVION   FLASR AGENGROCA     GPOKEPSAVILING SSSLIGGLERNOS, LEVIM     GSFAAP VCTSSTRMGVLKED     SEAR AGENGRAMAKAKK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GAPAGARGAKAMAKAK VGARRKAS     GALHKAT VGALHE KAMAKA     GALHKAT VGALHA GARGAK AL KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GALHKAT VGALHA KAMAKA     GA	1		l	1			FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
FLASRPAETIFLQEMYGLERSENPPDUGQLEGE   FRASPPCTSPRMQVLKED	1			(			PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
NLVPRIKRI,GIARVYKNPGQENGKOCGE, GPOKPDSAVYLINGSSLLGGRRNOSI,PVIM GSPAAPVCTSSPKMGVIKED		1				ļ	FLASRPAETFLOEMYGLRPSRNPPDVGOLKM
GPQKPDSA-VYLNSGSSLIGGERRNQSLFVIM GSPAAPVCTSSPKMGVLKED  A 5254 2 2621 LSLFGSRALGESGARAMAKAKVGARRKFO GAPAPVCTSSPKMGVLKED  A 5254 2 2621 LSLFGSRALGESGARAMAKAKVGARRKFO GAPAGAGGGARANSNFPEVKVNRQKCPOLG RKTRHDVGLPGVSRARALRKRTOTLLKE-YKG RDKSNVFRDKRFGSVNSNMSPEEKMMKRFA LEQQRHHEKKSTYNLNEDELITAHFGGSLADIE KHRDIVDSDSDAEDRGTLSGELTAAHFGGG GILHKKTQGGEGEREKYKSRKELELLAKSK QEKRERQAGREDALELTEKLDQDWKEIGTLAHFGGG GILHKKTQGGEGEREKYKSRKELELLAKSK QEKRERQAGREDALELTEKLDQDWKEIGTLAHFGGG GILHKKTQGGEGEREKYKSRKEHADDLNDGFV LDKDDRRLLSYKDGKMNVEEDVGEGSKEK SPESSNEEGEDSSGGDTESSDSPSHIDLES NVESEEENKRAKEORQTTGKGLISGKERAG KATRDELYTFAAPESYEERSLIGGSMEQ LLVVERIQKCNHPSLAGONKAKLEKLFGFLLE YGGLATDDPPI-TVDKLVVHLCGMFP ESASDAIKFVLRDAMHEMEMETKGRAALP GLDVINLKLTGLIFTSDFVHPVVTALGSMEP ALGKKNELLVVSAREDVATWQSSLSLRWA SRLRAFTSTEANHRLSCLAVGLALLKRCVHAL YGSLESFHAMGPIRALITHHADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPYPLK FTPRLVKVLEFGRKQGSSKEUERRDAE RKKKVQLFNSLATIGGEWKALKRKKKK RFKGAVERRKDNOPLARMQLESIMERDAE RKKKVQLFNSLATIGGEWKALKRKKKKK ASGSAGTIGVSHSCKQOSSKEEGERKRLHKHK REFKGAVERRKDNOPLARMQLESIMERDAE RKKKVQLFNSLATIGGEWKALKRKKKKK ASGSAGTIGVSHSCKSVEPVTYLSTLHAVRNVL FRRTPFLKSSFLSYDKEIPFILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTTCANVVHLPCTFPGWKIDGPAAAL EVLSSFFFFFFIKFSYKPONIV  692 2042 A 5282 56 1268 GMEPVCCCGECRGSSVDPRSTFVLSNLAEVV PMGSGSNRPQEIEGESGFALLFPQGGUKGIPGF HIKDDFNLTLERHUTLVWAFBEGREVVVIVYALTH FRYTTVLYMADSETFISLEECRGHKRAC KRKTSWUSGALAGAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWUSGALAGAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWUSGALAGAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWUSGALAGAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLTWADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLTWADSETFISLEECRGHKRAC KRKTSWEGALAGHLEGHLCVRVVAEEL ENVRIPHTVLTWAGGGEGDANTGGGLLK AGFPTNIVTLCHSIA  693 2043 A 5301 362 507 EERKERGGGULTVYVYFFIG				1		į	NLVDRLKRLGIARVVKNPGAOENGRCOFAFI
SSFAAPVCTSSPKMGVLKED	i	1	Ī	1			GPOKPDSAVYLNSGSSLLGGLRRNOSLPVIM
689 2039 A 5254 2 2621 S.S.F.GSRALGRSGARAMAKAKKVGARRKAS GAPAGARGGAKANSHPEVKVNROKPOLLO RKTRIDVGLPOVSRABALRKRTOTILLEYYE RDKSNVPROKEPGEYNSNNSPEEKMMKRPA LEQORHHEKKSIYNLINEDEELTHYYGQSLADIE KHNDIVDSDSDAEDGTLSGELTAAHFGGGG GLIKKKTQBEGEREKPKSKRELIEELIAKSK QEKRERQAQREDALLITEKLDQDWKEGITL SHKKTYRSSENGDKKEKPKPAYDMMYRELGF EMKAQPSNRMKTEALIAKEEQEHLRKLEAE RLRRMLGKDEDENVKKPKADDLINDGRV LDKDDRRLISYKDGKMMVEEDVQEEQSKEA SDPFSNEEGEGSSGGEDTEESPSSHLDLES NVESEGENEKPAKEGROTPGKGLISGKERAG KATRDELPYTTAAPESYVEELLIGRSMEEQ LLVVERIQKCNIPSLABGNKAKLEKLFGFILE YVGDLATIDPPDLTVIDKLVHYHLCQMFP ESASDAIKFVLRDAMHENGEBMETKGRALP GLDVLIYLKIGLIFTSTDFWTYTTAAVSV SRLRAFTSTEANHRILSCLAVGLALLKRCVLM YGSLSFHAMGPLRALLITHACSHOPLO ELQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRRQGSSEGERKLHKHK REFKGAVREIRKDNOFLARMQLSEBMERDAE RKRKVKQLTRSLATGGEGERKLHKHK REFKGAVREIRKDNOFLARMQLSEBMERDAE RKRKVKQLTRSLATGGEGERKLHKHK REFKGAVREIRKDNOFLARMQLSEBMERDAE RKRKVKQLTSLATGGAGERKLHKHK REFKGAVREIRKDNOFLARMQLSEBMERDAE RKRKVKQLTSLATGDFPTW ASQSAGTIOVSHCSWPVTVSLTLHAVRNVL FKITFFLKSSSFLSYDKEIFFILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVBLPGTPPGWKIDGPAAAL EVLSSFFFFLKSSSFLSYDKEIFFILIVLKPYLVTLT SFVK ASQSAGTIOVSHCSWPVTVSLTLHAVRNVL FKRTFPLKSSCALAGHLBEGHCLVRVVAEEL ENVRLIPHTVLYMADSSTFISLECCRGHRAR RKRSMETALALEKLFPKQCVITYFGIVVT PMGSGSNRPQEIEIGESGFALLFPGIBKIOFF HFKDPNLTLEHQLTEVGLDMPELRVVLV FGYNCCKVGASNYLQQVVSTSSIMMILAGQ QVDNLSSLTSEKSPHDIDDASCHSSFGHRI QSATVLLHEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVCAADAFR KFFFSVPLFGFGRGGIGCDRYTONFTLRKCN EVKDDDJHHSVTTIMALHLGSSK  693 2043 A 5301 362 507 EEIECRRFGFGGIGCDFTVTONFTLRKCN EVKDDDJHSVTTIMALHILGSSK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAMAAA		1	ļ	1	1	ļ	GSFAAPVCTSSPKMGVLKED
GAPAGARGPAKANSNPETVKVNRQKPQILG RKTRHDVGLPGVSRARALRKRTQTILLEYKE RDKSNVPRDKPGGPYNSNMSPEEKMMKRFA LEQQRHHEKLSTYNLINEDEELTHYGGGSLADE KHNDIVDSDSDAEDBGTLSGELTAAHFGGGG GLIHKKTQGEGEREKPKSRKELIEELIASK QEKKERQAQREDALELTEKLDQDWKEIQTTL SHKTPKSENRDKKEKPKPDAYDMMYRELGF EMKAQPSNRMKTEAELAKEGGHLRKLEAE RLRRMLGKDEDEVVKPKHMSADDLNDGFV LDKDDRRLLSYNDGKMVENGTALLOG LLVVERIQKCNHPSLAAGNGADDLDGFV LDKDDRRLLSYNDGKMVENGTALLOG RKATDELPYTPAAFESYEELSLLIGRSMEGE LLVVERIQKCNHPSLAAGNKAKLEKLFGFLE YVGOLATIDPPDLTVIDKLYHTHLCQMFP ESASDAIKFVLRDAMHEMEEMETKGRAALP GLDVLIYLKITGLIFTSDFWPWYTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRIFTELNFILGILYAPTNAHV GSLPSFHAIMGPLRALITDHLADCSFRHQELQ ELCOSTLTEMESQKCLCRELYCEKSKPVPLKL FTPRLVXVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRNDNQFLARNQLSEBMERDAE KRKVXQLINSLATQFGEGWKALKKKKKK KRKRKVKQLINSLATQFGEGWKALKKKKKKK ASGRAGITGVSHSWATVYLSTILHAVRNVL FFRTPLIVETGFHIHVGQAGLELLTSGDPFW ASGRAGITGVSHSWATVYLSTILHAVRNVL FKRTPFLKSSSTLSYDKEIPFILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTHCTANWVHLPGTPPGWKIDGPAAAL FYFFVFVETGFHHVGQAGLELLTSCDPFW ASGRAGITGVSHCSWPTVTVSTILHAVRNVL FKRTPFLKSSSTLSYDKEIPFILIVLKFYLVTLT SFVK  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSTVMSAGLAEAGHLECHLVRVVAEEL ENNRLIPHTVLYMADSSTFISLEECGHKRAR KRTSMETALALEKLFKQCQVT,GIVTPGIVVT PMGSGSNRPQEIEIGESGFALLPQUBGKQPF HFKDPNKLIT.EHQLTEVGLUDPFELRVLIV FGYNCKVAGASNYLQQVVSTFSSMMILAGG QVNNLSSLTSEKSPNPLIDIASCHSSGGRILL GGATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQVYRAGNVEADAFR KFFPSVPLFGFFGONGGICORVITGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK FFFSVPLFGFFGONGGICORVITGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK KTRAMRRLNMHEENLKTKKQHRERELYPL RKYAAAAA  694 2044 A 5310 1 204 RVLTARHTILKENLRFFYKGKKKKKKKKKR KTRAMRRLNMHEENLKTKKQHRERELLYPL RKYAAAAA	689	2039	A	5254	2	2621	
RKTRHDVGLFGVSRAALRKRTQTILLKEYKE RDKSNVFRDKFGFYNSNKFEMMKRFA LEQGRHEKKSIYNLINEDEELTHYGGGSLADIE KHNDIVDSDSDADBGTGTLSGELTAHGGGG GLLHKKTQQBGEEREKPKSRKELIEELLAKSK QEKRERQAQREDALELTEKLDQDWKELGYFLAYDMMYRELGF EMKAQPSNRMKTEALLAKEEQEHLRKLEAE RLRRMLGKDEDENVKKFKHSADDLINDGFV LDKDDRRLLSYKDGKMVEEDVQEEQSKEA SDFPSNEEEGGSSGGEDTEESDSPDSHLDLES NVESEENBERDAKEEQROTPGKGLISGKERAG KATRDELPYTAAPESYEERSLLLGRSMEEQ LLVVERIQKCNHPSLAEGNKAKLEKLEGFILLE YVGDLATDDPPDLTVIDKLVVHLLYCLGMFP ESASDAIKFVLRDAMHENMEBTIGRAALP GLDVITJKLTGLLFFTSDFWHPVVTTALVCL SQLITKCPILSJQDVVKGLFLEYVALS QRFFELINFLLGILYJATPNKASQGSTLVHFR ALGKNSELLVVSAREDVATPLAVCL SQLITKCPILSJQDVVKGLFCLEYVALS QRFFELINFLGILYJATPNKASQGSTLVHFR ALGKNSELLVVSAREDVATPLAVCL SQLITKCPILSJQDVVKGLFCLESKEVPILK FTFRLVKVLEFGREQGSSKEEGERKRLIHKHK REFKGAVREIKKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK FTFRLGAVREIKENDOMPLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFKK FFFFFFILVFSSSFLSYDKEIFFILIVLKFYLVTLT SFVK  690 2040 A 5261 1 304 FFFFFFILVFSSFSSYDKEIFFILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHITTHCTANWVHLPGTPPGWKALKRKKFK FKFFFFFLKFSYKPQNIV FKRTFPLKSSFSSFSYDKEIFFILVILKFYLVTLT SFVK  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV FRINSTYWISAGLABAGHLEGHCLVRVVAEEL ENVRILPHTIV,TWADSSTFISECRGHRARA KRTSMETALALEKLFPRQCQVLGIVTPGIVVT FMGSGSNRPQEIEIGESGFALLFPQIBGKQPF HFIKDPKNLTLERHQLTEVGLIDNPELRVVLV FGYNCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRII QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFGNGGIGCDRIVTGNFLRKCN EVENDDLHFISTYTTMALHLGELEGGHRARA KFFPSVPLFGFGNGGIGCDRIVTGNFLRKCN EVENDDLHFISTYTTMALHLGELEGGHRARA KFFPSVPLFGFGNGGIGCDRIVTGNFLRKCN EVENDDLHFISTYTTMALHLGESGGCHLLFREKKN KTRAMRRILNMHEENLKTKKQHRKERLYPL RVYAAXA  694 2044 A 5310 1 204 RVLTARHITIKENLEKFYKGKKDKFLDLRFK KTRAMRRILNMHEENLKTKKQHRKERLYPL RVYAAXA		1			[ ~	2021	GAPAGAPGGDAYANGADECAYANDOYFOR O
RDKSNYRDERGETYNSMSPEEKHYRQSNLADIE  LEQQRHIEKKSYNIN.NESELTHYRQSNLADIE  KHNDIVDSDSDAEDRGTI.SGELTAA.HEGGG GLIHKKTOGGERERKEYRSKELIEELLASYS QEKRERQAQREDALELTEKLDQDWKEIQTL SHKYTKSENROKKEKPKPDAYDMWYRELGF EMKAQPSNRMKTEALELKEGGEHLRKLEAE RLRRMLOKDEDENVKRYKHMSADDLDNGGV LDKDDRRLLSYNDGKMVPEDVQEEQSKEA SDFSNEEGDSSGGEDITEESDSPDSHIDLES NVESEEENEKPAKEQROPKGKIJSGKERAG KATADELPYTTAAPSSYELERSLLLGRSMEEQ LLVVERIQKONIPSLAEDIKAELH.FOFLIE YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP ESASDAIRFVLRDAMHEMBEMIETKGRAALP GLDVLIYLKITGLLFFTSDFWHPVYTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFFELINFILGILVATAMAGSGTLVHFPF ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAFSTEANHRISCLAVGGSTLVHFPF ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAFSTEANHRISCLAVGGSTLVHFPF ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAFSTEANHRISCLAVGGSTLVHFPF ALGKNSELLVVSAREDVATWQQSSLSLRWA REFRGAVREIRESQQLCRFLYCCEKSKYPPLKL FTFRLVKVLEFGRQQSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCCEKSKYPPLKL FTFRLVKVLEFGRQQSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCCEKSKYPPLKL FTFRLVKVLEFGRQQSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTFRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTFRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCRFLYCEKSKYPPLVL FTRLVKVLEFGRQGSSKEGERRI.HHKHK REFRGAVREIRESQQLCCHCUSTSTILHAVRNVL FKRTFPLKSSSFLSYDMSTFVLSILLAVVL FKRTFPLKSSSFLSYDMSTFVLSILLAVVL FKRTFPLKSSSFLSYDMSTFVLSILLAVVL FKRTFPLKSSSFLSYDMSTFVLSILLAVVL FKRTFPLKPSSVKPOND' FMRDAGABALLTLAFCCVCVLGVTCVTGVTCHTGIVVT FMGSGSNRPQEIEGSGFALLPPOIEGICOP HTRIGFMFACVGCGEGRGSSVDPSTFVLSILLAGA QSATVLLHEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGGYDYARGNVCAADAFR KFFPSVPLFGFGONGGICODRIVTGNFLRKCN FKYDDDLHFISTTILLECGRIHKRA KRTSMETALALEKLFPRQCVLGISFGGRILL GGGATULTURDUSSEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGGYDYARGNVCAADAFR KFFPSVPLFGFFGONGGICODRIVTGNFLRKCN FYXDDDLHFISTTILLECGRIHKRCN FKYDDDLHFISTTILLECGRIHKRCN FKYDDDLHFISTTILLEGGRIHKAR KTSMETALALEKLFFYKGKKDKPLDLRFK KTRAMRRINMHEENLKTKKQHRERLLYPL RVA	1	ł	ł	l .	}		PETPUNCE POUR AR ALDERTON LESSEE
LEQORHHEKKSIYNLINDEDELITYGGSLADIE   KINDIVIDSDSDAEDRGTLSGELTAAHFGGGG    GILHKTQQEGEREKKKSKELIEELIAKSK    QEKZERQAQREDALELTEKLDQDWKEIQTIL    SHKTPKSENRDKKEKPKPAYDMMYRELGF    EMRAQFSNRMKTEAELAKEEQEHIRKLEAE    RIRRMLGKDEENVKKPKHMSADDLINDGFV    LDKDDRRLLSYKDGRMNVEEDVQEEQSKEA    SDPESNEEGDSSGGGETUSSGESTUDLES    NVESEENEKPAKEQROTPGKGLISGKERAG    KATROLEPYTTAAPESYEDSYDSHILDLES    NVESEENEKPAKEQROTPGKGLISGKERAG    KATROLEPYTTAAPESYEDSYDSHILDLES    NVESEENEKPAKEQROTPGKGLISGKERAG    KATROLEPYTTAAPESYEDSILDLESILLGRSMEEQ    LLVVERIQKCNHPSLAEGNKAKLEKLPGFLE    YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP    ESASDAIKFVLRDMHEMEMEITGRAALP    GLDVLIYLKITGLLPFTSDEWHPVVTPALVCL    SQLLTKCPILSLQDVVKGLFVCCFLEYVALS    QRFPELINFLLGILYVAREDVTPALVCL    QRFPELINFLLGILYVAREDVTPALVCL    QRFPELINFLLGILYVARAPGOSSLSLRWA    SRLRAPTSTEANHRILSCLAVGLALLKRCVLM    VGSLPSTHAMGPLRALLDLADCSHPQELQ    ELCOSTLTEMESQKQLCRPLTCEKSKPVPLKL    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    ASSAGITOVSHCSSYLEQEWKALKKKFKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    ASSAGITOVSHCSSYLEQEWKALKKKFKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    ASSAGITOVSHCSSYLEQEWKALKKKFKK    FTPRVKVLEFGRKQGSSKEEQERKLIHKHKK    ASSAGITOVSHCSSYLEQEWKALKKKFKK    FTRFVKVLEFGFHHYQGSLSLINGDPTW    ASQSAGITOVSHCSSYLEQEWKALKKKFKK    FTRFVKVLEFGFHHYQGSLLINGDPTW    ASQSAGITOVSHCSSYLEQEWRALKKKFKK    FTRFVKVLEFGFHHYQGSLLINGDPTW    ASQSAGITOVSHCSSYLEJENDTW    ASQSAGITOVSHCSSYLEJENDTW    ASQSAGITOVSHCSSYLEJENDTW    ASQSAGITOVSHCSSYLEJENDTW    ASQSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASQSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSHCSSYLEJENDTW    ASSAGITOVSH		1	}		1		RAIRHDVOLPGVSKARALKKRIQILLKEYKE
KINDIVOSDSAEDRGTLSGELTAAIFIGGGG GILHKKTQGEGEREKKYSKRELIEELIAKSK QEKRERQAQREDALELTEKLDQDWKEIQTILI SHKTPKSENRDKKEKPKPAVDMWYRELGF EMKAQPSNRMKTEAELAKEEQEHLRKLEAE RLRRMLGKDEDENVKKPKHMSADDLNDGFV LDKDDRRLLSYKDGKMVEDEVQERQSKEA SDPESNEEGDSSGGEDTESDSSPSHLDLES NVESSEEREKPAKEQROTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMEEQ LLVVERIQKCNHPSLAEGNKAKLERLFGFLLE YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP ESASDAIKFVLRDAMHEMEEMETKGRAALP GLDVLIYLKITGLLFPTSDHPVVTPALVCL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS ORFIPELINFILGII VIATRASQGSTLVHPFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALKRCVLM YGSLPSSTHAMGPLRALLTDHLADCSHPQELQ ELQSTLTEMESQKQLCRFLTCEKSFVPLKL FTFRLVKVLEFGRKQGSSKEPGERKRLIHKHR REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKKERKEKKFKK REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKKALKRKKKKK REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKKALKRKKKKKK REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKKALKRKKKKKK REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKALKFRKKKFKK REFKGAVREIRKDNQFLARMALSEKKFKK REFKGAVREIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKALKFRKKFKK REFKGAVEIRKDNQFLARMQLSEEMERDAE RKRKVKQLFNSLATQEGKALKFRKKFKK REFKGAVEIRKDNQFLARMALSEKKFKK REFKGAVEIRKDNQFLARMALSEKKFKK REFKGAVEIRKDNQFLARMALSEKKFKK REFKGAVEIRKDNQFLARMALSEKKFKK REFKGAVEIRKDNQFLARMALSEKTRAK RETSMETALALEKLAVALEL ENVELPHTVLTVAMDSETPILVKFYLJNLAFVV PMGSGSNFRPQEIEGSGRALLFPQIEGKKIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVULV PGGSGNFRPQEIEGSGRALLFPQIEGKKIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVULV PGGSGSRPQEIEGGSGRALLFPQIEGKKIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVULV PGGSGSRPROGEIEGGCRATUTGNFLIKKCN EVYDDDLFHSYTIMALHHLGSSK FFPSVPLFGFFGGEIGCDRIVTGNFLIKKCN EVXDDDLFHSYTIMALHHLGSSK FFPSVPLFGFFGGEIGCDRIVTGNFLIKKCN EVXDDDLFHSYTIMALHHLGSSK FFPSVPLFGFFGGEIGCDRIVTGNFLIKKCN EVXDDDLFHSYTIMALHHLGSSK FFPSVPLFGFFGGUFTYWYGFIQELDCNRERGILLK ACPFTNIVTLCHSLA  694 2044 A 5310 1 204 RVLTAINHTLKENLKTKKQHKKERLYPL RKYAAKA	1	1		í		1	
GILHKKTQQEGEREKYKSRKELIEELIAKSK  GRERERQAGREDAL ELTLÖJÖVKEGITLI SHKTPKSENRÖKKEKPKPDAYDMMVRELGF EMKAQFSINRMKTEAELAKEGGEHIRKLEAE RIRRMLGKDEDENVKKPHMSADDLINDGFV LDKDDRRLLSYKDGRMNVEEDVQEEQSKEA SDPESNEEGGDSSGGGEHDSSPSSHLDLES NVESEENEKPAKEQROTPGKGLISGKERAG KATRDEDYTTAAPESYEELISLLGRSMEEQ LLVVERIQKONHPSLAEGNKAKLEKLFGFLE YVGDLATDDPPDLTVIDKLVYHLYHLQOMFP ESASDAIKFVLRDAMHEMEMIETKGRAALP GLDVLIVLKITGLLFPTSDFWHPVYPALVCL SQLITKCPHISLQDVVKGHVCCTFLEYVALS QRFPELINPLLGILYIATPNKASQGSTLVHPFR ALGKNSELLVVSAREDVATPQQSSLSLRWA SRRAPTSTEANHRIASLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHQELQ ELCQSTLTEMESGKQLORPLTCEKSKPYPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLHRHK REFKGAVREIRKDNQFLARMQLSEIMERDAE EKRKVKQLFNISLATGEGEWKALKKKKRK FTPRLVKVLEFGRKQGSSKEEQERKRLHRHK REFKGAVREIRKDNQFLARMQLSEIMERDAE FRREVINSTEAHHPUNGAGLELITSGDPFTW AGSSAGTICVSHCSWPVYLSTILHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFLKESYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCGEGCGSVDPSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISGALAEAGH EGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGTVTGIVVT PMGSGSNRPQCEIGGSGRALFPQIEGKGPF HFIKDPKNLTLERHQLTEVGLUNDSLRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGTVTGIVVT PMGSGSNRPQCEIGGSGRALFPQIEGRKQPF HFIKDPKNLTLERHQLTEVGLUNDSLRVVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGTVTGIVVT PMGSGSNRPQCEIGGGSGALLFPQIEGKKQPF HFIKDPKNLTLERHQLTEVGLUNDSLRVLV GGYATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRSPQTYTRAKGNVEADAPER KFPPSVPLFGFFGNGEIGCDRIVTGNFLRKCN EVKDDDLFHSTYTIMALHHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVTWYGFIQELDCNRERGILLK ACPPTNIVTLYHSTYTIMALHHLGSSK 694 2044 A 5310 1 204 RVLTAINHTLKENLKTKKQHRKERLYPL RKYAAKA	1	1	l	ł	ì	i	
ORKERGAGREDALELTEKLDODWKEIGITLE   SIKITYKSENROKKEKPKAYDNMMYRELGF   EMKAQPSNRMKTEAELAKEEQEHLRKLEAE   RIRRMIGKDEDENVKKPKHMSADDINDGGV   LDKDDRRILSYKDGKMKDEVDQEEQSKEA   SDPESNEEGDSSGGEDTEESDSPBEILDLES   NVESEEENEKPAKERGYTPGKGLISGKERAG   KATRDELPYTFAAPESYEELRSILLIGRIMEED   LLVVERIGKCHMESLAGEMKALEKLFGFILE   YVGDLATDDPPDLTVIDKLVVHLYHLCGMFP   ESASDAIKFVLRDAMHEMEMEITSGRAALP   GLDVLYKLIGKCHMESLAGEMEVTVPALVCI   SQLLTKCPILSLQDVVKGLFVCCLFLYVALS   QRFPELINFILGII, YAIFMKASGGSTLVHPR   ALGKNSELLVVSAREDVATWQOSSLSLRWA   STRAPTSTEAMHRILSCLFLYWAGAEDLATVHOPA   ALGKNSELLVVSAREDVATWQOSSLSLRWA   SGLRSTSTAMHRILSCLFSKVPLKL   FTFRLVKVLFGRKQGSSKEQERKRLIKKRCVLM   YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ   ELQSTLTEMESQKQLCRFLTCEKSKVPLKL   FTFRLVKVLFGRKQGSSKEQERKRLIKKRKK   REFKGAVREIRKDNQFLARMQLSEMERDAE   RKRVKVGLPNSLATGEGWKALKRKKFKK   REFKGAVREIRKDNQFLARMQLSEMERDAE   RKRVKVGLPNSLATGEGWKALKKKFKK   REFKGAVREIRKDNQFLARMQLSELTSGDPPTW  ASQSAGITGVSHCSWPVIYVISTILHAVRNVL   FKRTPILKSSSFLSYDKEIPFLIJVLKPYLVTLT   SFVK    FYRIFPLKSSSFLSYDKEIPFLIJVLKPYLVTLT   SFVK    GOMEPVGCGGEGGSSVDPRSTFVLSNLAEVV   ERVLTFLPAKALLRVACVCRLWRECVRRVLR   THRSVTWISAGLEAGHLEGHCLYVVAEEL   ENNRILPHTVL YMADSETTSILEECRGHKRAR   KRTSMETALALEKLIPKQCQVLGIPTGIVVT   PMGSGSNRPQEEIGESGFALLFPQIEGKIQPF   HTKLDPKNLTLERHQLTEVGLLDNPELRVVLV   FGYNCCK VGASNYLQVVCJSFSGHRI   QSATVLLNEDVSDEKTAEAMGRIKAANIPE   HNTIGFMFACVGRGFQYTRAMGNICADAPER   KFFPSVPLFGFFGNGEIGCDRIVTGNPLIKCN   EVKDDDLFHSYTIMALHILGSSK   693   2043   A 5301   362   507   EEIKERFGFGUTYWYGFIQELDCNRERGILLK   ACPPTNIVTLCHSIA   694   2044   A 5310   1   204   RVLTAINHTLKENLRKFKKQHKKERLYPL	1			]	ĺ		KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
SHKTPKSENRDKKEKPKPDA YDMAYNELIGF EMKAQPSNNKKTEALAKEEOGHLRKLEAE RILRRMI.GKDEDENVKKPKHMSADDLNDGFV LDKDDRRILSYKDGKMNVEEUFUVGEGVSKE SPESNEEGDSSGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSILLGRSMEEQ LLVVERIQKCHPSLAGONKAKLEKLPGFILE YVGDLATDDPPDLTVIDKLVHLLYHLCVMF BSASDAIKFVLRDAMHEMBEMETKGRAALP GLDVLIYLKITGLIFFTSDFWIHVVTPALVCL SQLITKCPILSQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSSKEEQERKRLIKHK REFKGAVERIKNDJOATTOQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLFSHAIMGPLRALLTDHLADCSHPOEL ELCQSTLTEMESQKQLCRPLTCEKSKFVPLKL FTPRLVKVLEFGRRQGSSKEEQERKRLIKHKK REFKGAVERIKNDJOATMOJSEBMERDAE RRKVKQLFNSLATOEGEWKALKRKKFKK FFFFVLVETGFRHVGQAGLELLTSGDPPTW ASQSAGTIGVSHCSWPVTYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIPPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTT-CTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNTV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV ERVLTT-LPAKALLRVACVCRLWRECVRRVLR THSSVTWISAGLAEGEGFGALLFPQIEGIKUPP HFIKDFKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQVASTESDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAMMCRLAAMPE HNTIGFMFACVGRGFQYTRAKGNVEADAFR KFFSVPLEGFFGNGEIGCDRIVTGNFTLRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGGGLVTYWYGFFQELDCNRERGILLK ACFFTINTVITCHAL				1	l	!	GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
SHKTPKSENRDKKEKPKPDA YDMAYNELIGF EMKAQPSNNKKTEALAKEEOGHLRKLEAE RILRRMI.GKDEDENVKKPKHMSADDLNDGFV LDKDDRRILSYKDGKMNVEEUFUVGEGVSKE SPESNEEGDSSGEDTEESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSILLGRSMEEQ LLVVERIQKCHPSLAGONKAKLEKLPGFILE YVGDLATDDPPDLTVIDKLVHLLYHLCVMF BSASDAIKFVLRDAMHEMBEMETKGRAALP GLDVLIYLKITGLIFFTSDFWIHVVTPALVCL SQLITKCPILSQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYIATPNKASQGSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSTLJHFPF ALGKNESLLVVSARGOSSKEEQERKRLIKHK REFKGAVERIKNDJOATTOQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLFSHAIMGPLRALLTDHLADCSHPOEL ELCQSTLTEMESQKQLCRPLTCEKSKFVPLKL FTPRLVKVLEFGRRQGSSKEEQERKRLIKHKK REFKGAVERIKNDJOATMOJSEBMERDAE RRKVKQLFNSLATOEGEWKALKRKKFKK FFFFVLVETGFRHVGQAGLELLTSGDPPTW ASQSAGTIGVSHCSWPVTYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIPPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTT-CTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNTV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV ERVLTT-LPAKALLRVACVCRLWRECVRRVLR THSSVTWISAGLAEGEGFGALLFPQIEGIKUPP HFIKDFKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQVASTESDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAMMCRLAAMPE HNTIGFMFACVGRGFQYTRAKGNVEADAFR KFFSVPLEGFFGNGEIGCDRIVTGNFTLRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGGGLVTYWYGFFQELDCNRERGILLK ACFFTINTVITCHAL	1	1	1		1		QEKRERQAQREDALELTEKLDQDWKEIQTLL
EMKAQPSNRMKTEAELAKEEQEHLRKILEAE   RIRRMIGKDEDENVEKPHMSADDLDIDGFV     LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA     SDPESNEEGBSSGGDTEESDSPDSHLDLES     NVESEEBNEKPAKEQRQTPGKGLISGKERAG     KATRDELPYTFAAPESYEERSILLGRSMEEQ     LLVVERIGKCNIPPLSAGNKAKLEKLPGFILE     YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP     ESASDAIKFVLRDAMHEMEEMHETKGRAALP     GLDVLTYLKITGLLFYSDFWHEVVTPALVCL     SQLITKCPILSLQDVVKGLFVCCLFLEYVALS     QRFPELLENFLIGILYLTHKASQGSTLVHPFR     ALGKNSELLVVSAREDVATWQQSSLSLRWA     SRIRAPTSTEANHRLSCLAVGLALLKRCVLM     YGSLESTHAMIGPLASGKEQDEKRLIHKHK     REFKGAVREIRKDNQFLARMQLSEIMERDAE     RKRKVKQLENSLATOGEEWKALKKKKFKK     FFRFVLVVLEFGRFHIPGQGGLELLTSGDPFTW     ASQSAGITGVSHCSWPVTYVLSTLLHAVRNVL     FRRTPFLKSSFLSYDKEIFPILVLKFYLVLTT     SPVK     691   2041   A   5270   3   158   NCHITHCTANWHLPGTPFGWKIDGPAAAL     EVLSSFFFLKFSKYPONIV     692   2042   A   5282   56   1268   GMEPVGCCGCRGSSVDPSSTFVLSNLAEVV     ERVLTFLPAKALLRVACVCR LWRECVRRVL     THRSVTWISAGLAEAGHLEGHCLVRVVAEEL     ENVRIPHTVL YMADSETFISLECRGHKRAR     KRTSMETALALEKLFPKQCQVLGIVTPGIVT     PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF     HFIKDPKNLTLERHQLTEVGLLDPPELRVVIV     FGYNCCKVGASNYLQQVVSTFSDMNILAGG     QVDNLSSLTSEKNPLDDASGVVGLSFSGHRI     QSATVLLADENDSELTAAAMQRLKAANIPE     HNTIGFMFACVGRGFQYYRAKGNVEADAFR     KPFPSVPLFGFFGNGEIGCDRIVTGNFLRKCN     VKDDDLIPHSTTITMALHLGSSK     693   2043   A   5301   362   507   EEIKERFGFGLVTYWGFJELDCNRERGILLK     ACFPTINVTLCHSIA     ACFPTINVTLCHSIA     A   COMPANIAL     COMPANIAL	1		İ			]	SHKTPKSENRDKKEKPKPDAYDMMVRELGF
RILRRMI.GKDEDENVKKPKHMSADDL.NDGPU LDKDDRRLLSYKDGENVEEQSKEA SDPESNEEGDSSGGEDTESDSPDSHLDLES NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTTAAPSYSELRSILLCRSMEQ LLVVERIQKCNIPSLAGONKAKLEKLFOFLLE YVGDLATDPPDLTVLVILVLLCOMFP ESASDAIKFVLRDAMHENEEMEETKGRAALP GLDVLIYLKITGLIFFTSDFWIPVVTPALVCL SQLITKCPILSLQDVVKGLFVCCLFLEYVALS QRFIPELINFLLGILYJATPINKASQGSTLVHFPR ALGKNSELLVVSARGDVATWQQSSISLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKFVPLKL FTPRLVKVLEFGRQGSSKEEQERKRLIHKHK REFKGAVREIRKNDFQLARMQLSEIMERDAE RKRKVKQLFNSLATOGGEWKALKRKKFKK RFKGAVREIRKNDFARMQLSEIMERDAE RKRKVKQLFNSLATOGGEWKALKRKKFKK FFFYFLVETGFFHVGQAGLELTSGDPFTW ASQSAGITGVSHCSWPVITYLSTLLHAVRNVL FKRTFPLKSSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKDGPAAAL EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV ERVLTTLPAKALLRVACVCRL WRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVAFEL ENVRLPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTGVVT PMGSGSNRPQEIEIGESGFALLFPQEIGIKQPF HFIKDFKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQVSTFSDMMILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTARAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVFADAFR KFFPSVPLGFFGNGEIGCDRIVTGNFLRKCN EVKDDDLHFSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGRGLVTYWYGFIQELDCNRERGILLK ACFFINIVITLCHIAL 694 2044 A 5310 1 204 RVLTANHTLKENLKFYKGKKKKPKLYPL	J	1		1		f	EMKAQPSNRMKTEAELAKEEOEHLRKLEAE
LDKDDRRILSYKDGKMNVEEDVOGEGSKEA SPPESNEEGGDSGEDSPDSDPHDLDES NVESEENEKPAKEQRQTPGKGLISGKERAG KATTADELPYTAAPESYEELRSLLIGRSMEQ LLVVERIQKCNIPISLAEGNKAKLEKLIGFILE YVGDLATDDPPDLTVIDKLVVHLYHLCOMFP ESASDAIKFVLRAMHEMEEMIETKGRAALP GLDVLIYLKITGLIFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLFLGYVALS QRFPELINFLLGILSQTAVGNGSSTLVHFFR ALGKNSELLVVSAREDVATTWQOSSIS.SLRWA SRLRAPTSTEANHIRLSCLAVGIALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPOELQ ELCQSTLITEMESQKQLCRPLTCEKSKPVPLKL FFPRLVKVLEFGRKQGSSKEPQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRVKQLFNSLATQEGEWKALKRKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE ASQSAGITOVSHCSWPUTVLSTILLHAVRNVL FRTPFLKSSSFLSYDKEIPFILVLKFYLVTLT SPVK  690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELITSGDPPTW ASQSAGITOVSHCSWPUTVLSTILLHAVRNVL FRYTPLKSSSFLSYDKEIPFILVLKFYLVTLT SPVK  691 2041 A 5270 3 158 NCHITHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKRSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPSTFVLSNLAEVV ERVLTEPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALERHPQCQVVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKLOPF HFIKDPKNLTLERHQLTEVGLUDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSBCATAAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLLGFFFGNOEIGCDRIVTGNFLRKCN GSATVLLNEDVSBCATAAAMGRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLLFGFFGNOEIGCDRIVTGNFLRKCN EVKDDDLFHSYTITMALHLICSSK  693 2043 A 5301 362 507 EEIKERFGPGLVTWWYGFIGLEDCNRERGILLK ACFPITIVITLCHSL 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKFDLRPK KTRAMRRRLIMHEENLKTKKQHRKERLYPL RKYLAMARRLIMMHEENLKTKKQHRKERLYPL RKYLAMARRLIMMHEENLKTKKQHRKERLYPL	1	1	ì	1		ľ	RLRRMLGKDEDENVKKPKHMSADDLNDGFV
SPESNEEEGDSSGGDTESDSPDSHLDLES NVESEEENERPAKEQROTPOKGLISGKERAG KATRDELPYTFAAPESYEELRSLLIGRSMEEQ LLVVERIQKCNHPSLAGSNKAKLEKLFGFLE YVGDLATDDPPDLTVIDKLVYHLYHLCOMFP ESASDAIKFVLRDAMHEMEEMETKGRAALP GLDVLIYLKITGLIFTSDFWHPVVTPALVCL SQLITKCPHSLQDVVKGLFVCCLFLEYYALS QRFFELINFLIGHJATPINKASQGSTLVHFR ALGKNSELLVVSAEDVATWQOSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGFWKALKRKKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGFWKALKRKKFKK ASGAGTIOVSHCSKPVPVLSTLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKPYLVTLT SFVK  691 2041 A 5270 3 158 NCHITHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFLKPSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGCGCGSVDPTSTFVLSNLAEVV ERVLITFLPAKALLRAVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMOESTFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGTVTPGIVVT PMGSGSNRPQEIEIGESGFALFPQEGIKLOPF HFIKDPKNLTLERFLQLTEVGLLDINGELRVALV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMORLKAANIPE HNTIGFMFACVGRGFQYYRAGNVEADAFR KFFSVPLFGFFGNGEIGCDRIVTGNFLRKCN EVKDDDLFHSYTTINALHLGSSK  693 2043 A 5301 362 507 EEIKERRGPGLVTYWYGFTQELDCNRERGILLK ACFPTINVTLCHSS  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKFLDLRPK KTRAMRRRLINMHEENLKTKKQHRKERLYPL RKYAAKA	1	1		;		!	LDKDDRRLLSYKDGKMNVEEDVOEEOSKEA
NVESEEENEKPAKEQRQTPGKGLISGKERAG KATRDELPYTFAAPESYEELRSLLLGRSMEEQ LLVVERIQKCNIHPSLAEGNKAKLEKLEGFILE YVGDLATDDPPDLTVIDKLVYHLYHLCQMFP ESASDAIKFVLRADAMHEMEEMETKGRAALP GLDVLIYLKITGLIFPTSDFWHPVVTPALVCL SQLLTKCPILSLQDVVKGLIFVYALS QRFPELINFLLGIL YIATPNKASQGSTLVHFPR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLSFSHAMGPIRALLTDHLADCSHQPELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK FFFFLWETGFHHVGQAGLELLTSGDPPTW ASQSAGITOVSHCSWPVIVVLSTLLHAVRNVL FKRTPPLKSSSFLSYDKEIFPILVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSIFIFFILKSSYRPQNV  692 2042 A 5282 56 1268 GMEPVGCCGECGSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECGHKARA KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALFPQIEGIKIQPF HFIKDPKNLTLERFIQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMVIILAGG QVDNLSSLTSEKRPQCIFFLRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERRGPGLVTWYGFIQELDCNRERGILK ACFPTINVTLCHSL 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKFLDLRPK KTRAMRRRLIMHEENLKTKKQHRKERLYPL RKYAAKA	1	1	1				SDPESNEEEGDSSGGEDTEESDSPDSHI DI ES
KATRDELPYTFAAPESYEERSLLIGRSMEEQ   LLVVERIOKCHPISLAEGNKAKLEKLEGILE   YVGDLATDDPPDLTVIDKLVYHLYCMFP   ESASDAIKFVLRDAMHEMEEMIETKGRAALP   GLDVLIVLKITGLIFPISDFMPVVTPALVCL   SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS   QRFFELIFILLGILYIATPKASQGSLSLIWA   ALGKNSELLVVSAREDVATWQQSSLSLIWA   SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM   YGSLPSFHAIMGPLAALLTDHLADCSHPQELQ   ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL   FTPRLVKVLEFGRQGSSKEEQERKRLIHKHK   REFKGAVREIRKDNOPLARMQLSEIMERDAE   RKRVKQLFNSLATQEGEWKALKRKKFKK   FTFRLVKVLEFGRQGSGSKEEQERKRLIHKHK   REFKGAVREIRKDNOPLARMQLSEIMERDAE   RKRVKQLFNSLATQEGEWKALKRKKFKK   FTFFVLVETGFHHVGQAGLELLTSGDPTW   ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL   FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT   SFVK   SF	1	1	l	1 1			
LLVVERIQKCNIPSLAEGNKAKLEKLFGFLE   YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP   ESASDAIKFVLRDAMHEMEEMETKGRAALP   GLDVLIVLKITGLLFFTSDFWHPVVTPALVCL   SQLITKCPILSIVAMHEMEEMETKGRAALP   GLDVLIVLKITGLLFFTSDFWHPVVTPALVCL   SQLITKCPILSIVAMOVYKGLFVCCIFLEFYVALS   QRFIPELINFLLGILYIATPNKASQGSTLVHFR   ALGKNSELLVVSAREDVATWQQSSLSLRWA   SRLRAPTSTEAHHRLSCLAVGLALLKRCVLM   YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ   ELCQSTLTEMESQKQLCRRITCEKSKPVPLKL   FTPRLVKVLEFGRKQGSKEEGERKRLIHKHK   REFKGAVREIRKDNOFLARMQLSEIMERDAE   RKRKVKQLFNSLATQEGEWKALKKFKK   REFKGAVEIRKDNOFLARMQLSEIMERDAE   RKRKVKQLFNSLATQEGEWKALKKFKK   ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL   FKRTFPLKSSSFLSYDGEGWKALKKFKK   FTPRLVKVLEFGTHIVGQAGLELLTSGDPPTW   ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL   FKRTFPLKSSSFLSYDGEGWKALKKFKK   FVFRV   FVFRV   FKRTFPLKSSSFLSYDGEGWKALKKFKK   EVLSSFFFFLKFSYKPQNIV   FKRTFPLKSSSFLSYDGEGKGREFTLIVLKFYLVTLT   SFVK   STVK   SFVK   GMEPVGCCGERGSSVDPRSTFVLSNLAEVV   ERVLTFLPAKALLRVACVCRUWRECVRRVLR   THRSVTWISAGLAEAGHLEGHCLVRVVAEL   ENVRLPHTVLYMADSETTSLEECRGHKRAR   KRTSMETALALEKLFPKQCQVLGIVTPGFVVT   PMGSGSNRPQEGEGGBGFALLFPQDEGKQPF   HFIKDPKNLTLERQLTEVGLUPSERVVLV   FGYNCCKVGSNYLQCVVSTFSDMVILAGG   QVDNLSSLTSEKPPLDIDASGVVGLSFFGHRI   QSATVLLNEDVSDEKTAEAAMQRLKAANIPE   HNTIGFMFACVGRGFQYYTAKGNVEADAER   KFFSPVPLFFFGFGGEGGCDRIVTONFILRKCN   EVKDDDLFHSVTTIMALHLGSSK   EEIKERFFGFGIGCGDRIVTONFILRKCN   EVKDDDLFHSVTTIMALHLGSSK   EEIKERFFGFGICTUTVYVGFTQLKCN   EVKDDDLFHSVTTIMALHLGSSK   EEIKERFFGFGIVTVWYGFTQELDCNRERGILLK   ACCPPINIVTLCHSIA   ACCPPINIVTLCHSIA   ACCPPINIVTLCHSIA   ACCPPINIVTLCHSIA   KTAMMRRINNHEENLKTKKQHRKERLYPL   RVYAAKA	j	j	)				KATRDEI DVTEA ADEQVEET DOLLI CROLEGO
YVGDLATDPPPDLTVIDKLVVHLYHLCOMFP   ESASDAIKFVLRDAMHEMEEMIETKGRAALP   GLDVLIVLKTIGLIFFTSDFWHPVVTPALVCL   SQLLTKCPILSLQDVVKGLFVCCLFLETYALS   QRFIPELINFLLGILYIATPINKASQGSTLVHFFR   ALGKINSELLVVSAREDVATWQQSSLSLRWA   SRLRAPTSTEANHRILSCLAVGLALLKRCVLM   YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ   ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL   FTPRLVKVLEFGRKQGSSKEQERKRLHKKHK   REFKGAVREIRKDNQFLARMQLSEIMERDAE   RKKVKQLINSLATQEGEWKALKRKKFKK   FTPRLVKVLEFGRHVGQAGLELLTSGDPFTW   ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL   FKRTFPLKSSSLSYDKEIPFLIVLKPYLVTLT   SFVK   NCHTTHCTANWVHLPGTPPGWKIDGPAAAL   EVLSSFFFFLKFSYKPQNIV     691   2041   A   5270   3   158   NCHTTHCTANWVHLPGTPPGWKIDGPAAAL   EVLSSFFFFLKFSYKPQNIV   ERVLTSFFFFLKFSYKPQNIV   ERVLTFLPAKALLRVACVCRLWRECVRVLR   ERVLTFLPAKALLRVACVCRLWRECVRVVLR   ERVLTFLPAKALLRVACVCRLWRECVRVVLR   ERVLTFLPAKALLRVACVCRLWRECVRVVLR   ENVRILPHTVLYMADSETFISLEECRGHKRAR   KRTSMETALALEKLFPKQCQVLGIVTPGIVVT   PMGSGSNRPQELIGESGFALLPQUEGKIQFF   HFIKDPKNLTLERHQLTEVGLLDMPELRVVLV   FGYNCCKVGASNYLQQVVSTFSDMNILAGG   QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI   QSATVLLNEDVSDEKTARAAMQRLKAANIPE   HNTIGFMFACVGRGFQYYRAKGNVEADAFR   KFFPSVPLGFFGNGEIGCDRIVTGNFILRKCN   EVKDDDLFHSVTTIMALHLGSSK   EKKERFGPGVYYRAKGNVEADAFR   KFFPSVPLGFFGNGEIGCDRIVTGNFILRKCN   EVKDDDLFHSVTTIMALHLGSSK   EEIKERFGPGVYTVYGFIQELDCNRERGILK   ACCPPTNIVTLCHSIA   A   5310   1   204   RVLTANHTILKENLRFYKGKKDKPLDLRPK   KTRAMRRLNNHEENLKTKKQHRKERLYPL   RVYAAKA							I I WEDION CHIDGE ARCHITANT BY BOTH TO
ESASDAIKFVLRDAMHEMEEMIETKGRAALP GLDVLIYLKITGLIFPTSDFWHPVVTPALVCL SQLLTKCPILSQDVVKGLFVCCLFLEYVALS QRFPELINFLLGILVATPINKASQGSTLVHFFR ALGKNSELLVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALITDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSKEEQERKRLHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVTYVLSTLLHAVRNVL FKRTPPLKSSSFLSYDKEIPPILIVLKFYLVTLT SFVK NCHTTHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFFLKFSYKPONIV 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFFLKFSYKPONIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFJAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEEIGGSGFALFPQEGIKQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAF KFFPSVPLFGFFGNGEIGCDRIVTGNFLRKCN EVKDDDLHISYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGFGLVIYWGFTQELDCNRERGILK ACFPTINIVILCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			1				VVCDI ATOODOO TUDVI VARIATION COL
GLDVLIYLKITGLLFPTSDFWHPVVTPALVCIL SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFPELINFLGILYIATPNKASQGSTLVHFFR ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHILADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVTYVLSTLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPFGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDFRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRI-WRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADESTFISLECRGHKRAR KRTSMETALEKLFPKQCQVLGIVTPGIVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQFF HFIKDFKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGANTJQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRCN EVKDDDLHSYTTIMALHILGSSK 693 2043 A 5301 362 507 EEIKERFFGGLVIVWGFIQELDCNRERGILLK 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRLNMHEENLKTTKKQHRKERLYPL RKYAAKA	J	1	ļ	j .			T VGDLATDDPPDLT VIDKLY VHLYHLCOMFP
SQLTKCPILSLQDVVKGLFVCCLFLEYVALS QRFPELINFLLGILYIATPNKASQGSTLVHPFR ALGKNSELVVSAREDVATWQOSSLSLRWA SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPKLK FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGWKALKRKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTILHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHITHCTANWHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRLPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTFGIVVT PMGSGSNRPQBIEIGESGFALLFPQIEGIKIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTITMALHLGSSK 693 2043 A 5301 362 507 EEIKERFOFGLVTYWYGFIQELDCNRERGILLK ACFPTINVILCHSIA 694 2044 A 5310 1 204 RVLTANHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	1						ESASDAIKF VLRDAMHEMEEMIETK GRAALP
QRFIPELINFLLGIL YIATPNKASQGSTL VIIFFR ALGKNSELLVYSAREDVATWQQSSLSRWA SRIRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPFTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSVKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTTLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTGFIVTFOFVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQVYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTITMALHILGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFJQELDCNRERGILLK ACFPINIVILCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	1			1 1			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
ALGKNSELLVVSAREDVATWQQSSLSLRWA SRLAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPS;HAIMGPLRALLITDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPFLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALFPQEIEIGESGFALFPVEISCHIQF HFIKDPKLTLERHQLTEVGLLDNDELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDLFHSYTIMALHHLGSSK 693 2043 A 5301 362 507 EEIKERFGFGLVIYWGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	}			1 1			SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKKKFKK 690 2040 A 5261 1 304 FFFFVPLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKPYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGFGLVIYWYGFJQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTANHITLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	1 -			1 1			QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
FOR STATE AND ST	1	1		1 1			ALGKNSELLVVSAREDVATWQQSSLSLRWA
FOR STATE AND ST	ł	1		!!!	1		SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHR REFKGAVREIRKDNQFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVITVLSTLLHAVRNVL FKRIFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQEGIKIOFF HFIKDPKNLTLERHQLTEVGLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMVIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHILGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIVWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVILTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL							YGSLPSFHAIMGPLRALLTDHLADCSHPOELO
FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKKFKK FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVIR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRIPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACPPTNIVTLCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	ſ	1 1		1			ELCOSTLTEMESOKOLCRPLTCEKSKPVPLKI.
REFKGAVREIRKDNOFLARMQLSEIMERDAE RKRKVKQLFNSLATQEGEWKALKRKFFKK  690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACPPTNIVTLCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRLNMHEENLKTKKQHRKERLYPL RKYAAKA							FTPRLVKVLEFGRKOGSSKFFOFRKRI THKHK
RKRKVQLFNSLATQEGEWKALKRKKFKK 690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELLTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK 691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFFLKFSYKPQNIV 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRVVLRE THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEAADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIVWYGFIQELDCNRERGILLK ACFFINIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA							REFEGAVREIRKDNOEL ARMOL SEIMERDAE
690 2040 A 5261 1 304 FFFFVFLVETGFHHVGQAGLELTSGDPPTW ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCGGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHILGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA  694 2044 A 5310 1 204 RVLTANHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL		1 1					
ASQSAGITOVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTITHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCGGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQFF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIVWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL	690	2040	Δ	5261	1	204	
FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT SFVK  691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	""		A	1201	• •	JU4	
SFVK	1	(					
691 2041 A 5270 3 158 NCHTTHCTANWVHLPGTPPGWKIDGPAAAL 692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFFSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	1				l		
EVLSSFFFFLKFSYKPQNIV  692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	-	2041					
692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIVWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	ז לם	2041	Α [	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
692 2042 A 5282 56 1268 GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA							EVLSSFFFFFLKFSYKPQNIV
ERVLTFLPAKALLRVACVCRLWRECVRRVLR THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHITVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	692	2042	Α	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
THRSVTWISAGLAEAGHLEGHCLVRVVAEEL ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			- 1				ERVLTFLPAKALLRVACVCRLWRECVRRVI R
ENVRILPHTVLYMADSETFISLEECRGHKRAR KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	}		l		- 1	l	THRSVTWISAGLAEAGHT.EGHCT.VBVVAFET
KRTSMETALALEKLFPKQCQVLGIVTPGIVVT PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			J		j		ENVRILPHTVI.YMADSETEISI FECEGLIKD AD
PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			1		[	Ī	KRTSMETALALERI EDROCOVI CINTROTATE
HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK 693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			1		i	l	PMGGGGND PODIDIGEOGRAFI E PROPERTIES
FGYNCCKVGASNYLQQVVSTFSDMNIILAGG QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA					]	]	
QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI QSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			- 1	1	1		
GSATVLLNEDVSDEKTAEAAMQRLKAANIPE HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			ł	l	1	ł	
HNTIGFMFACVGRGFQYYRAKGNVEADAFR KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVILCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA		ļ	j	1	j		QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN EVKDDDLFHSYTTIMALIHLGSSK  693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA  694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA				į			
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA		}	1	1	ł	l	
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			J	J	J	į	KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
693 2043 A 5301 362 507 EEIKERFGPGLVIYWYGFIQELDCNRERGILLK ACFPTNIVTLCHSIA 694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA			[			-	
694 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA	693	2043	Α	5301	362	507	
594 2044 A 5310 1 204 RVLTAINHTLKENLRKFYKGKKDKPLDLRPK KTRAMRRRLNMHEENLKTKKQHRKERLYPL RKYAAKA							
KYETAINTI EKENEKAT TROKADAPEDERPK KTRAMRRELNMHEENLKTKKQHRKERLYPL RKYAAKA	694	2044	A	5310	1	204	
RKYAAKA			- (		-		KTR AMPRILAMMEENT VTVVOITSVENT INT
	1		)		1	Į.	
695 2045 A 5315 125 1596 ETRSTAVKSEVOVCISITICLED TYPEKAYD	695	2045	<del>_  </del>	5315	125	1596	
695 2045 A 5315 125 1596 ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP			Α	נוננ	147	1370	EIRSIAVKSEVQVCISLLLCLEDRTMPKKAKP

NO. of much coids sequence where the continuation of the continuat	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucleotide seq unce clo							
uence benee							
Sequence   9/495							
uence    914   ng to first amino acid residue of peptide pepti			ĺ		f -		
amino acid residue of peptide residue of peptide sequence   TaThreonine, V=Valine, W=Typtophan, Y=Typtophan, Y=Typtophan, Y=Typtophan, Y=Dyptophan, Y=Typtophan, Y=Dyptophan, Y=Typtophan, Y=Typtophan, Y=Dyptide sequence   TaThreonine, V=Valine, W=Typtophan, Y=Typt		1					
Persidue of   peptide   sequence   y=Tyrosine, x=Unknown, =Sing podon,   peptide insertion   molecutide insertion   molecutide insertion   molecutide insertion   ToSGKEEGPAPCKQMKLEAAGGPSALNFDSP   SSLFESLSPIKETFREEWEGKPLLIQRIDPPA   LATYYGSLEKLTDLIKSLSKGMYYGDVNV   CRCVNGKKKVLNKDGKAIFLQLRKDFDQKK   ATIOFIQPORKNDE   WRIGHELECTFGSLVGS   NYVITPAGSQG1PPHYDDVEVFLQLEGEKH   WRLY3PTVPLAREFYSVAEAEDIGRPVHEPML   KRGDLLYFPRGTHQADTPAGLAHSTHVTIST   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTAKEDVBLATT   YQNNSWGGPLLDTISGLVFDTATT   YQNNSWGGPLLDTISGLT   YQNNSKETTHMANGESTEFFT   GGKLPRLISHLDALKQTWNSPAISVKDLKLTTDE   KESELJSLWTECLDQVV   YEARLANGSEGWTKYSTLATT   YQNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPLATT   YGNNSWGGPTAT   YGNNSWGGTAT   YGNNSWGGPTAT		}		1			
Poptide   Popt	1	[					
	ł	l				Sequence	
1056KEEGPAPCKQMKLEAAGGPSALNFDSP   SSLFESLEPKETEFKEEVECKPLLIQRIDPA   LATYYGSLFKLITDLKSLCSIGMYYGBLVNV   CRCVNGKKVLNAKDGKAHFLQLKKDFDOXK   ATTOFFGPQPRKDEL WRIQEKLECYFGSLVGS   NYUTFAGSQLLPPHODVEVFULQEGEKH   WRLYSPIPVPLAREVSVEAEEUGRPVHEPML   KFGDLLYPFRGTHIAODTPAGLALSTHWTIST   YONNSWGDFLLDTISCLYPDTAKEDVELRTG   EFROLLLOYESTIVATRRLSGFLRTALADRLEG   TRELLSSDMKKDPHMRLPPYSAGDGALSTP   GGKLPKLDSVYRLQFKDHVLTVLPDQCQDS   ETQEKMYYTISLKNSKTHMMGNEETEFH   GLRPFLSILDALAKQIWNSPAISVKDLKLITDE   EKSELVALSWELGLYST   GGKLPKLDSVYRLQFKDHVLTVLPDQCQDS   ETQEKMYYTISLKNSKTHMMGNEETEFH   GLRPFLSILDALAKQIWNSPAISVKDLKLITDE   EKSELVALSWELGLYST   EKSELVALSWELGLYST   EKSELVALSWELGLYST   EKSELVALSWELGLYST   EKSELVALSWELGLYST   EKSELVALSWELGLISDITVKLLRSSSQTIET   SLQDIDSSLSPGGSLADAWAHQEGTTPKDRN   VEKLQVLLNTENSYQTKEKNSEEWIKMHLRKQLLSLTNQCPD   EEEVSKYQSTVNLQELTPLQMATASGICHT   MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE   YEARLNSEEWIKMHLRKQLLSLTNQCPD   EEEVSKYQSTVNLQELTPLQMATASGICHT   MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSLTNDGGRNVDCL   LARNHILBESGGSTYCTLDGGST   LARNYTGSLFRANGE   VSPSWGWSRTDFR   LARNHILBESGGST   LARNYTGSLFRANGE   LARNHILBESGGST   LARNYTGSLFRANGE   LARNHILBESGGST   LARNYTGSLFRANGE   LARNHILBESGGST   LARNYTGSLFRANGE   LARNHILBESGGST   LARNYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   LARNHILBESGGST   LARNSYTGSLFRANGE   L	1		i				
SSLFESILSPIKTETFKEFWEGKPLLIGRIDPPA   LATYYGSLFKLTDILSLESKGMYYGDNVV   CRCVNGKKYUNKDGKAHFLQLRKDFDQKK   ATIOFHQPORKDEL WRIGKLECYFGSLVGS   NYVITPAGSQG1PPHYDDVEVPLQLEGEKH   WRLYSPITVPLAEVSVAEREDIGRYVHEPML   KPGDLLYFFRGTHIOADTPAGLAHSTHVTIST   YQNNSWGGPLLDTISCLYPTAKEDVERTG   FRQLLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FRQLLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FRQLLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FRQLLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLST   FROLLQYESTTVATRRLSGFLRTLADRLEG   TKELLSSDMKROFHMRILPYSAGDOALLSSGGTFT   SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN   FRESKYLGYTYNELQETLYPKKMARSGGITT   FROLLDYMKFTEATFGLYFT   FROLLDYMKFTAKSGEN   FRONKRISEWINKHHLRKQLLST.INGCFD   FEEVSKYQEYTNELQETLYPKKMTASSGIKHT   FRYSTRATH   FRY	<b></b>				Sequence	<del></del>	
LATTYGSLFRLTDLKSLCSRGMYYGRDYOK	Į.						
CRCVNGKKKVINKDGKAHFLQLRDFDQKS ATIGPHOPORYEDLWRIGEKLECYTOSLVGS		1					
ATIOFHOPOREKDELWRIGERECTYGSLVG   NVPITTPAGSGGIPPHYDDVEVELQLGERKH     WRLYPPTVPIAREYSVBAGERIGRPVHERNI     KPGDLLYPPERTHIOADTRIGLAHSTHVTIST     YQNMSWGDFLLDTISGLVPDTAKEDVELRTIG     FRQLLLQVESTTVATRLSGFLRTLADRLEG     TRELLSSDMKKDFIMHRLPPYSAGDOAELSTP     GGKLPRLDSVVRLQRKDHVLTVLPDQDQSD     TERLSSDMKKDFIMHRLPPYSAGDOAELSTP     GGKLPRLDSVVRLQRKDHVLTVLPDQDQSD     TERLSSDMKKDFIMHRLPPYSAGDOAELSTP     GGKLPRLDSVVRLQRKDHVLTVLPDQDQSD     GGKLPRLDSVVRLQRKDHVLTVLPDQDQSD     TRELLSSDMKKDFIMHRQHESTEFH     GLRPFLSHLDALKQIWNSPAISVKDLKLTTDE     EESHLVLSLWTECLIQVV     LMKXYLEAAELGEISDIHTKLLRLSSSQGTIE     LODDSRLSPGGLADAWAHQEGTHPKDRN     VEKLQVILINGMTSITYQFKKDKAERRLAYN     VEKLQVILINGMTSITYQFKKDKAERRLAYN     TERLHRDAKLKEEMEGVVKE     LEANNHLESGGSLADAWAHQEGTHPKDRN     VEKLQVILINGMTSITYQFKKDKAERRLAYN     THE PROPERTY	1						
NYJTTPAGSGGLPPHYDDVEVELQLEGEKH   WRLYPHYPIARSYSVAGERIGRYVEFML   KPGDLLYPFRGTHIQADTFAGLAHSTHVTIS   YQNNSWGDFLLDTSGLVPDTAGLAHSTHVTIS   TYQNNSWGDFLLDTSGLVPDTAGLAHSTHVTIS   GKLFRLDSVRLQFKDHVLTPDQDQSTD   ETQEKMYJYHSLKNSRETHMMGNEGETEFH   GLRPISHLDALKQHKSHNISHLTPDQDQSTD   ETQEKMYJYHSLKNSRETHMMGNEGETEFH   GLRPISHLDALKQHWSPATSVKLLTPDQDQSTD   ETQEKMYJYHSLKNSRETHMMGNEGETEFH   GLRPISHLDALKQHWSPATSVKLLTDDQCSTD   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EKESLVLSLWTECLIQVV   EQHIKYDKQKLVYHATKAMTHFTDECVKK   YSAPLNSSEGWIRKMHLRKQLSTNOCTOL   EEFVSKVQEYTNELQETLYQKMFTASSGIKHT   MTHYPSSTILVEMTLGMKKLKEEMEGVVKE   LAENNHILBSGGSLDMGGLRVVCL   LAENNHILBSGGSLDMGGGRVVEDCSLUFF   EFFVSKVQEYTNELQETLYQKMFTASSGIKHT   MTHYPSSTILVEMTLLGMKKLKEEMEGVVKE   LAENNHILBSGGSLDMGGGRVVEDCSLUFF   EFFVSKVQEYTNELQETLYQKMFTASSGIKHT   MTHYPSSTILVEMTLLGMKKLKEEMEGVVKE   LAENNHILBSGGSLDMGGGRVVCL   LAENNHILBSGGSLDMGGGRUKTHVCH   EFFVSKVGESTAGREGALHEL   KLEEWKGRLQVTEHLPEKISSLQEDEPENDA   KKEALLNLPRNPSVUDKQDKD   KKEALLNLPRNPSVUDKQDKD   EFFVSKTVGTKYGKQKKKALHEL   KLEEWKGRLQVTEHLPEKISSLQEDEPENDA   KKEALLNLPRNPSVUDKQDKD   EFFVSKTVGTKYGKACKKALHEL   KLEEWKGRLQVTEHLPEKISSLQEDEPENDA   KKEALLNLPRNPSVUDKQDKD   EFFVSKTVGTKYGKACKKALHEL   KLEEWKGRLQVTEHLPEKISSLQEDEPENDA   KKEALLNLPRNPSVUDKQDKD   EFFVSKTVGTKYGKACKKALHEL   EFFXVSKTSKALHEL   EFFXVSKTVGTKYGALQCKM   YLYMFKLLTJELEKYXVTCHALQCKM   YLYMFKLLTJELEKYXVTCHALQCKM   YLYMFKLLTJELEKYXVTCHALQCKM   YLYMFKLLTJELEKYXVTCHALQCKM   YLYMFKLTJEDRKTVALAQCKM   EFFXVSKTVGTKYGACKACKKALLUTHANGER   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCENTOTTHOS   EFFXVSKTVALAQCE				İ	İ		
WRLYHPTYPIAREYSYLAEERIGRPVIEFPAL   RFGDILTYPRGTIIIQADTBGGLASTHIVTIST   YQNNSWGDFLIDTISGLYPDTAKEDYELRTG   PROLLIQVESTTYATRILSGFLATLADRIEG   TKEILSSDMKKDFIMHRLPYSAGDOADLSTD   GKEPLDSVPLGKCDHIVTYRRISGFLATLADRIEG   TKEILSSDMKKDFIMHRLPYSAGDOADLSTD   GKEPLDSVPLGKCDHIVTYPOADSD   GKERLDSVVRLQKCDHIVTYDEDASD   GKEPLDSVRLQKCDHIVTYDEDASD   GKERLDSVRLQKCDHIVTLATESSQCTIET   GKEPLSSVRLAEIGEISDHTKLIRLSSSQCTIET   LANGENDRIFTYGRKDKAERRLAYN   EEQHKPDXQKLYVHATKAMTHTDECVKK   YEAFLNKSEEWIRKIMHLRIGSGTHYKDRN   VEKLQVILINGMTBIYYQFKDKAERRLAYN   EEQHKPDXQKLYVHATKAMTHTDECVKK   YEAFLNKSEEWIRKIMHLRIQLISTINCOFT   EESYSKYQEYTHELGETLDKOKHPTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   MTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   WTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   WTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   WTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   WTPIYPSSTILVENTLGETLDKOKHFTASSGIKHT   WTGILLANGAGTASKFAA   GFGABSVPUETVOKOKGKIKTITVCPFFIKTGELLAUTHALSTILVENTLAAT   GKERNACHFTASSGIKHT   WTGILLANGAGTASKFAA   GFGABSVPUETVOKOKGKIKTITVCPFFIKTGELLAUTHALSTILVENTLAAT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIKTASSGIT   GKERNACHFTASSGIT   GKERNACHFTASSGIT   GKERNACHFTASSGIT   GKERNACHFTASSGIT   GKE			ł	1			
KPGDLLYPPRGTHPQAGLAISTHVTIST	1			İ		ŧ	
YQNNSWGDFILDITISGLVFDTAKEDVELRTG   PRQLILQVESTVATRALSGLEGO   FRQULIQVESTVATRALSQLEGO   FRQUILQVESTVATRALSQLEGO   FTQEKMVYLYHSLKNSRETHMMGNEESTEFH   GRFFLSHLDALKQIWNSFAISVKDLKLTTDE   EKESLVISLWTELIQVV   EKESLVISLWTELIQVV   EKESLVISLWTELIQVV   LMKCYYLEAAELGEISDHTKLLKLISSGOTIET   SLQDDSRLSPGSSLADAWAHGEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQLYYHATKAMHOEGTHFKDRN   VEKLQVLINCMTEIYYOFKKDKAERRLAYN   EGOHKEPKOKQHYHALGSICHYHATKASGKHT   MTPTYPSSNTLVEMTIGLWSKKEEMEGVYK   VERAFIN   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMDGILRNVDCL   LARNNHILESGGSLTMVGLAYCASGC   VSSSWPGWSFTPDR   LPERKSLRSVRSGPFTSQSPITRNLDGTAGSGC   LARNHILESGGSLTMVGLAYCASGC   LARNHILESGGSLTMVGLAYCASGC   LARNHILESGGSLTMVGLAYCASGC   LARNHILESGGSLTMVGLAYCASGC   LARNHILESGGSTONGCLADYCASKFAA   FOFASSVPVETRYQKXGKITVCPFFIKTGM   FEGCTTGCPSLLPILEPKYAVEKIVEALIQEEM   YLYMPKLLYFMMFLKSFLAKGSWFGNPTSLDKEE   GPENFRORGSGGGISVSTTLIGGFSRUFKKV   VETQAQULSTHDQPSVQLADEKNGAQTTP   AGAPRSLQPPGRPDFELSSSEG   PEPSPRDGSGGGGISVSTTLIGGFSRUFKKV   VETQAQULSTHDQPSVQLADEKNGAQTTP   AGAPRSLQPPGRPDFELSSSEG   FEPSPRDGSGGGISVSTTLIGGFSRUFKKV   VETQAQULSTHDQPSVQLADEKNGAQTTP   AGAPRSLQPPGRPDFELSSFELAKGSWFGNPTICKEEL   GSPSVPCTGCEHFFVAEQDKLLKHMITEHIV   LADVKLVADFRJVLYWKRFTEQPITDFCSV   RINSTAPFEQENYFLLODVLPDRIRLEELQ   KQRREBLEQQQQENDINTHIONCOMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAQCQCENDINTHIOCCMFCNEEF   CTQKKLDILAWAGAAA   CTGKKLDINTA   CTGKKLDINTA   CT	ł i	1		i			
PRQLILQVESTTVATRRLSGEARTLADREG   TEELLSSDMKKDFIMHRLPYSAGDOABLSTP   GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD   ETQEKMVYIYHSLKNSKETHMMGNEEBIETHFI   GLRFPLSHLDALKQINNSPAISVKDLKLTTDE   ERESLVLSLWTECLIQVV	1			ł			
TEELLSSDMKKDFIMIRLPPYSAGDGAELSTP	1			]			
GGKLPRLDSVVRLQPKDHIVLTVLPQDQQSD				l			
696 2046 A 5318 1476 742 LMKXYLEAABLGEISDIHTKLIRLSSSGÖTIET 696 EKESLVLSLWTECLIQVV  LMKXYLEAABLGEISDIHTKLIRLSSSGÖTIET 697 2047 A 5320 244 478 LOYNFLYENGENTOGLANWAHQEGTHPKORN 698 2048 A 5324 266 714 LPIRKSERWIRKHLHIRKQLLSLTNQCFDI 698 2048 A 5324 266 714 LPIRKSLESVRSGFTSQWHDLGSLQPP 698 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA 698 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA 698 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA 698 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA 698 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA 698 2049 A 5344 3 614 PTAEEMSLLTPENFVDROKOK 699 2049 A 5344 3 614 PTAEEMSLLTPENFVDROKOK 699 2049 A 5344 3 614 PTAEEMSLTPENFVDROKOK 699 2050 A 5344 3 614 PTAEEMSLTPENFVDROKOK 699 2050 A 5344 3 614 PTAEEMSLTPESSPLAKRSWFGNFISLDKE 699 2050 A 5344 3 614 PTAEEMSLTPESSPLAKRSWFGNFISLDKE 699 2050 C 614 PTAEEMSLTPESSPLAKRSWFGNFISLDKE 699 2050 C 700 2		l					
GLRPFLSHLDALKQIWNSPAISVEDLKLTTDE							
	î i			ĺ		ĺ	
696						•	
SLQDIDSRLSPOGSLADA WAHOGGTHPKDEN VEKLQVLLNCMTEIYYQFKKDKABERLAYN EEQHKFDKQKLYYHATKAMTHFTDECVKK YEARLNSEE WIRKMLHLRXQLISLTINQCFDI EEEVSK VQEYTNELQELTLPQKMFTASSGIKHT MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHLESGGSLTMDGGLRNVDCL  697 2047 A 5320 244 478 LDYNFFLFENTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDPE  698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAPGVXSGEFVERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKQDG  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGGTTIGCPSLLPILERKYAVEKIVEALLQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSFELAKRSWFONFISLDKEE QIFLVLKDKPLSSIKADDYHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKVRFQVDISSSEG PEFSPRROGSGGGIYSVTFTLISGPSRFFKRV VETIQAQLLSTHIDQPSVQALADEKNGAQTIP AGAPPRSLOPPFGRPDESSSPRROFPKRDV KELATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLLEPLSSPRGGFTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFGNTLWKRFFTEQHTDPCSV RINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRNSVUNHMAREHAFNIGLPDNTVNCNFEL CTLQKLLDNLQCLYCEKTFRDKNTLKDHMR KKQHRLNPKNREVPGRYDCVMFTSK KQHRKNRKNREVPGRYDCVMCNEFL LGRNSVUNHMAREHAFNIGLPDNTVNCNFEL CTLQKLLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKNRKNREVPGVARQHQCRCYGCTYGCTWKFKS KADLRTHMEEITKHTSILLPDRKTWDQLEYFF TYENDTLLWTLSDSSDLTAQEQNENVPIISE DTSKLYALKQSSLINALLL							
VEKLQVILINCMTEIYYQFKKDİXABRRLAYN EEQİHKFDKQKLYYHATKAMTIHFTDECVK YEAFLINISEE WIRKMİLIRKQLLSLITIQÇFDİ EEEVSKYQEYTINELQETLIQKIMFTASSIĞİKİİ MTPİYPİSSINLVEMTİLĞİKKİLKEEMEĞÜYVKE LAENNHILESĞİLTINDĞĞİRNIYDÇL 697 2047 A 5320 244 478 LDYNFFLFEMTFĞÜYSQAĞVQWHDLĞSLQPP PPĞFKQFSCLSLPSSWDYRHLPPİLLANFSREĞ VSPSWPĞWSRTPDFR 698 2048 A 5324 266 714 LPİRKSLİSVİSĞİNTÜNLÜĞTASĞİĞ LAKTVTĞİLFRINVĞLRĞÜVAĞĞİİĞALLĞTP VĞĞİLMAPÇKYSĞETIVÇERKÇİNDİKADİĞİ KLEEWKĞİLQİVTEHLİPKLESSİĞİLĞİPFINDA KKÜBALINLPRIPSIVDİKQİKDİ CÖŞÜLMAPÇKYSĞETIVÇERKÇİNDAKALHEL KLEEWKĞİLQİVTEHLİPKLESSİĞİÇİFFINDA KKÜBALINLPRIPSIVDİKQİKDİ TYMPĞİLLİYİNMİLİSİŞİŞİKİĞİLİDÜYLĞİ LHAMDĞFADÇİKİ 700 2050 A 5344 3 614 PTAEEMSİLTPİSŞİFLAKİŞİVFKİĞLİLDƏVLĞİ LHAMDĞFADÇİKİ 701 2051 A 5346 3 1383 HASVLPĞİYBİLİŞİŞİŞİSİSİNİLİŞİ ÇİŞİRÇİNĞİNİLİŞİİLİŞİŞİŞİŞİĞİĞİ ÇİŞİŞİRÇİĞĞİĞİYİTİLİĞĞİŞİŞİFŞİŞĞİTİLE ĞİŞİYÇİĞİÇİĞİLİŞİLİŞİŞİŞİŞİĞİÇİ ÇİŞİŞİŞİŞİŞİŞİİLİŞİLİŞİŞİŞİŞİĞİİ ÇİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞİŞ	696	2046	Α	5318	1476	742	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
EEQHKPOKQKLYYHATKAMTHFTDECVKK YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI EEEVSKYQEYTNELQETLPQKMFTASSGIKHT MTPTYPSSNTLVEMTLGMKKLKEEMEGVYVE LAENNHILESGGSLTMDGGLRNYDCL  697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQPSCLSLPSSWDYRHLPPHLANFSREG VSSWPGWSRTPDFR  698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGGAESVFVEITVQKQKGIKTTIVCPFPIKTGM FEGCTTGCPSLLPILEPKYAVEKIVAALQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK YLYMPKLLSTHDQVSVQALADEKNGAGTRP AGAPPRSLQPPPGPPDELSSSPRGFYCDK UTIQAQLISTHDQPSVQALADEKNGAGTRP AGAPPRSLQPPPGPPDELSSSPRGFYCDK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLIEFLSLPESPGTTTLE GSPSVPCIFCEEHFPVAEQDKLKHMIEHKIV LADVKLVADFQRYLLYWKRFFTEGPTDFCSV RNISTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHOCMFCNEEF LGRRSVININMAREHAFNIGLPDNIVNCNEEL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREVFTYVINVTELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNIFREFYVINTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNIFREFYVINTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLTYTHMEDAHEFDLLKKISELG LNFYQQVKLVINTREGTFYTYVNTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLTYTHMEDAHEFDLLKKISELG LNFYQQVKLVINTREGTFYTYVNTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLTYTHMEDAHEFDLLKKISELG LNFYQQVKLVINTREGTFYTYVNTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLTYTHMEDAHEFDLLKKISELG LNFYQQVKLVINTREGTFYTYVNTVELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLTYTHMEDAHEFDLLKKISELG LNFYQQVKLVINTREGTFYTYVNTVELGKSWEE VGLEDDREGTTHLWTLSDESSDLTAQEQNENVPIISE DTSKLYALKQSSLNQLLL							SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
VEÁFLNKSEÉWIRKMI.HLRKQLISLTNQCFD    EEEVSKYQEYTNELQETLPQKMFTASSGIKHT   MTPIYPSSNTLVEMTLGMKKLKEMEGVVKE   LAENNHILESGGSLTMDGGLRNVDCL   LDYNFFLEMTFGLVSQAGVQWHDLGSLQPP   PPGFKQPSCLSLPSWDYRHLPPHLANFSREG   VSPSWPGWSRTPDFR   VSPSWPGWSRTPDFR   LPIRKSLRSVRSGPFTSQSPITRNLDGTASGSC   LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP   VGGLLMAFQKYSGETVQERCKQDRKALHEL   KLEEWKGRLQVTEHJPEKIESSLQEDFENDA   KKIEALLNLPRNPSVIDKQDKD   KKIEALLNLPRNPSVIDKQDKD   RPHGHLVCISSAGLSGVNGLADYCASKFAA   FGFAESVFVETFVQKQKGKGITTIVCPFFIKTGM   FGGCTTGCPSILPILEPKYAVEKIVEAILQEKM   YLYMPKILYFMMFLKSFLPLKTGLLIADFKM   FGGTTGCPSILPILEPKYAVEKIVEAILQEKM   YLYMPKILYFMMFLKSFLPLKTGLLIADFKE   QTFLVLKDKPLSSKAGGSVGVALADEKNGAOTRP   AGAPPR SLQPPPGRPDPELSSSPRRGPKDKK   LLATNGTPL   GSPSVPCITVQQLLSTHDQPSVQALADEKNGAOTRP   AGAPPR SLQPPPGRPDPELSSSPRRGPKDKK   LLATNGTPL   GSPSVPCITCEHFPVAEQDKLLKHMIJEHKIV   LADVKLVAADFQRYLIYMKRFTTEQPITDFCSV   RINSTAPFEEQENYFLLCDVLPEDRILREELQ   KQRLREILGQQQCBCNTFOTTLE   GSPSVPCITCEHFPVAEQDKLLKHMIJEHKIV   LADVKLVAADFQRYLIYMKRFTTEQPITDFCSV   RINSTAPFEEQENYFLLCDVLPEDRILREELQ   KQRLREILGQQQCENTYFILTICHTERNTLKDHMR   KKQHRKJNPKNREVPRYVINYLELGKSWEE   VQLEDDRELIDHQEDDWSDWEEHPASAVCL   FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG   LNFYQQVKLNNFRRQVHQCRCYGCRYCYGCNYCFKS   KADLRTHMEETKHTISLLPDRKTWDQLEYYFP   TYENDTILLWTLSDSESDLTAQEQNENVPIISE   DTSKLYALKQSSILNQLLL	1			1			VEKLQVLLNCMTEIYYQFKKDKAERRLAYN
EEEVSKYQEYTNELQETLPQKMFTASSĞIKHT MTPTYPSSNTLVEMTLGMKKLKEEMEGVVKE LAENNHILESGGSLTMDGGLRNVDCL LAENNHILESGGSLTMDGGLRNVDCL  697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPFKCKPSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR  698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERXQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRINPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGYNGLADYCASKFAA FGFAESVPVETFVQKQKGIKTTIVCPPFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEALLQEKM YLYMFKLLYFMMTLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSISKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGGGTYSVTFTLISGPSRFKKV VETIQAQLLSTHDQPSVQALADEKNGAQTTP AGAPPRSLQPPFGRPPELSSSPRGPFKDK  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLKHMIEHKIV LADVKLVADFGRYJLVWKRFFTEGPIDFCSV RINSTAPFEEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHCVCMFCNEEF LGRRSVILNHMAREHAFNIGLPDNTVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREVPKYVIVMEDAHERDLLKKSELG LNFYQQVKLVNTRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTISLLPDRKTWDQLEYYPF TYENDTILLWTLSDSESDLTARQEQNENVPIISE	[			[			EEQIHKFDKQKLYYHATKAMTHFTDECVKK
MTPTYPSSNTLVEMTI.GMKKI.KEEMEGVYKE LAENNHILESGGSLTMDGGLRNVDCL  697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR  698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRRNVGLRGLVAGGIGALLGTP VGGLIMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKODKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADVCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFPIKTGM FEGCTTGCPSLLPILEFKYAVEKIVEALQEKM YLYMPKILLYFMMFLKSFLPLKTGLILADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PPESPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDFELSSSPRRGPPKDKK LLATNOTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLLEPISLPESPGGTTTILE GSPSVPCFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPTEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQGERNDTNFHIGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKFRKRNTLDHMR KKQHRKINFKNRFYDRFYVNYLELGKSWEE VQLEDDRELLDHQEDWSDWEEHFBASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL				•			YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
MTPTYPSSNTLVEMTILGMKKLKEEMEGVYKE LAENNHLESGGSLTMDGGLRNVDCL  697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG VSPSWPGWSRTPDFR  698 2048 A 5324 266 714 LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC LAKTVTGSLFRRNVGLRGLVAGGIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKODKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADVCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFPIKTGM FEGCTTGCPSLLPILEFKYAVEKIVEALQEKM YLYMPKLLYPMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PPESPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNOTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLLEPISLPESGGTTTILE GSPSVPCFCEEHFPVAEQDKLLKIMHEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPTEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQGERNDTNFHIGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKFRKRNTLDHMR KKQHRKINFKNRFYDRFYVNYLELGKSWEE VQLEDDRELLDHQEDWSDWEEHFBASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYAALKQSSILNQLLL							EEEVSKYOEYTNELOETLPOKMFTASSGIKHT
697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPHLANFSREG VSPSWPGWSRTDFR 698 2048 A 5324 266 714 LPERKSLRSVRSGFPTSQSPITRILDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTFESSPELAKRSWFGNFISLDKEE QFFVLKDKFLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKFVRFQVDISSSEG PEPSPRRDGSGGGGTSSVTFTLISGFSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDELSSSPRGPFDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEFLSLPESPGGTTTLE GSPSVPCIFCEHFPVAEQDKLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAFFEQENYFLICDVLPEDRILREELQ KORLEILEQQQQERNDTINFHGVCMFCNEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFROKNITLKDHMR KKQHKNIPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTISLIPDRKTWDQLEYYFP TYENDTLLWTLSDESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
697 2047 A 5320 244 478 LDYNFFLFEMTFGLVSQAGVQWHDLGSLQPP PPGFKQFSCLSLPSSWDYRHLPHLANFSREG VSPSWPGWSRTDFR 698 2048 A 5324 266 714 LPERKSLRSVRSGFPTSQSPITRILDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD 699 2049 A 5334 699 277 RPHGHLVCISSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK 700 2050 A 5344 3 614 PTAEEMSSLTFESSPELAKRSWFGNFISLDKEE QFFVLKDKFLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKFVRFQVDISSSEG PEPSPRRDGSGGGGTSSVTFTLISGFSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDELSSSPRGPFDKK LLATNGTPL 701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEFLSLPESPGGTTTLE GSPSVPCIFCEHFPVAEQDKLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV RINSTAFFEQENYFLICDVLPEDRILREELQ KORLEILEQQQQERNDTINFHGVCMFCNEF LGRRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFROKNITLKDHMR KKQHKNIPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTISLIPDRKTWDQLEYYFP TYENDTLLWTLSDESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					:		- ·
698 2048 A 5324 266 714 EPRKSLRSYRSGPPTSQSPITRNLDGTASGSC LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP VGGLLMAFQKYSGETTVGERKQKDRKALHEL KLEEWKGRLQVTEHLPFKIESSLQEDEPENDA KKIEALINLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVPVETFVQKQKGKITTVCPFPIKTGM FEGCTTGCPSLLPILEPKKYAVEKIVEAILQEKM Y1.YMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSIKAJOVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPRSLQPPPGRPDPELSSSPRRGPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTGAVARMQEDRDGSC STVGGVGYGDSCDCLEPLSLPESSPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKFTEQPITDFCSV IRINSTAPFEGEONYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDITNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFREDKTITKCHMR KKQHRKINPKNREYDRFVVNYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTISLLPDRKTWDQLEYYFP TYENDTILLWTLSDSSSDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLD	697	2047	Α	5320	244	478	
VSPSWPGWSRTPDFR			• •	1000			
698 2048 A 5324 266 714 LPIRKSLRSVRSGFFTSQSPITRNLDGTASGSC LAKTVTGSLFRINVOLRGLVAGGIIGALLGTP VGGLMAFQKYSGETVQERQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKEALINLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYPMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEVKASGGPSVPQKPVRFQVDISSSEG PEFSPRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP-AGAPPRSLQPPPGRPDPELSSSPRRGPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STYGGVGGDSKDCLLEFLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYLJWRKRFTEQPITDFCSV IRNSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDDLCLVCVEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFVVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLIPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLIL			ľ				
LAKTVTGSLFRINVGLRGLVAGGIIGALLGTT VGGLLMAFQKYSGGTVQERCQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVPVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGFSVPQKPVRFQVDISSSEG PEPSPRRDSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYIL VWRKRFTEQPITDFCSV RINSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTIQKKLDNLQCLYCEKTFRDKNILKDHMR KKQFIRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKISELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	698	2048	Δ-	5324	266	714	
VGGLLMAFQKYSGETVQERKQKDRKALHEL KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALLNLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGYNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSLIPLEPKYAVEKIVEAILQEKM YLYMPKILYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHEDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTILE GSPSVPVCFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV REINSTAPFEEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGRNSSVILNHMAREHAFNIGLPDNIVNCNEEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKTNPKNREYDRRYVNYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDUTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	000	2040	Γ	3324	200	/14	
KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA KKIEALINLPRNPSVIDKQDKD  699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADVCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSILIPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVFQVDISSSEG PESPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCLLEFLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRVILYWRKRFTEQPITDFCSV IRNSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDITNFHGVCMFCNEEF LGNRSVILNIHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILIPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGETTGCPSILPILEPKYAVEKIVEAILQEKM YLYMPKILYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLISTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGDSKDCILEPLSLPESPGGTTILE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLVCEKTRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDREILDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMETKHTSLLPDRKTWQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		,					
699 2049 A 5334 699 277 RPHGHLVCISSSAGLSGVNGLADYCASKFAA FGFAESVPVETFVQKQKGIKTTIVCPFPIKTGM FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKILLYPMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGTYSVTIFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV LADVKLVADDFRYILYWRKRFTEQPITDFCSV IRNISTAPFEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDREILDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHYKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	ŀ		l				
FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM FEGCTTGCPSILPILEPKYAVEKIVEAILQEKM YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLOPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDREILDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	600	2040	_	5224	600	222	
FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM YLYMPKILYFMMFILKSFLPLKTGLLIADYLGI LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCTLEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV LADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNTVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVNYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFTRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	099	2049	Α .	2224	099	2//	
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQFSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
LHAMDGFADQKK  700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENVFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTILQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSILPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
700 2050 A 5344 3 614 PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLILL	1					i	
QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRDGSGGGGIYSVIFTLISGPSRFFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTILLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG PEPSPRRDGSGGGGIYSVTFTLISGPSRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEHFPVAEQDKLLKHMIEHKIV IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	700	2050	Α	5344	3	614	
PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1						
VETIQAQLLSTHDQPSVQALADEKNGAQTRP AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRNSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
AGAPPRSLQPPPGRPDPELSSSPRRGPPKĎKK LLATNGTPL  701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1.						
TOTAL TOTAL TRANSPORT TO THE PROPERTY OF THE P	[						VETIQAQLLSTHDQPSVQALADEKNGAQTRP
701 2051 A 5346 3 1383 HASVLFCRVMAASKTQGAVARMQEDRDGSC STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQGERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL					•		LLATNGTPL
STVGGVGYGDSKDCILEPLSLPESPGGTTTLE GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	701	2051	Α	5346	3	1383	HASVLFCRVMAASKTQGAVARMQEDRDGSC
GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV IADVKLVADFQRYILYWRKRFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL						•	
IADVKLVADFQRYILYWRKFFTEQPITDFCSV IRINSTAPFEEQENYFLLCDVLPEDRILREELQ KQRLREILEQQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
IRINSTAPFEQÈNYFLLCDVLPEDRILREELQ KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
KQRLREILEQQQERNDTNFHGVCMFCNEEF LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL	1 1						
KKQHRKINPKNREYDRFYVINYLELGKSWEE VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
VQLEDDRELLDHQEDDWSDWEEHPASAVCL FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		l					
FCEKQAETIEKLYVHMEDAHEFDLLKIKSELG LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
KADLRTHMEETKHTSLLPDRKTWDQLEYYFP TYENDTLLWTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL		.0					
TYENDTLL WTLSDSESDLTAQEQNENVPIISE DTSKLYALKQSSILNQLLL							
DTSKLYALKQSSILNQLLL							· · · · · · · · · · · · · · · · · · ·
702   ZUS2   A   5356   ZSUZ   1540   MAAATRGCRPWGSLLGLUGLVSAAAAAWD	L	20.55		50.57		1510	
	/02	2052	Α	5356	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalenine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	İ	09/496	соггеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		<b>!</b>		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
<b>\</b>		İ		peptide		/=possible nucleotide deletion, \=possible
ļ				sequence		nucleotide insertion
		]				LASLRCTLGAFCECDFRPDLPGLECDLAQHL
		[			•	AGQHLAKALVVKALKAFVRDPAPTKPLVLSL
1		}				HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH
1		ł				HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA
1		J			j	CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV
					1	VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI
					į	MEERLLDAVVPFLPLQRHHVRHCVLNELAQL
		i			}	GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK
					ŀ	TVASRIAFFL
703	2053	A	5380	278	657	LFLOKLRMKTEEEARTHTEIEMFLRKEOOKL
'**	2000		3500	2.0	1 037	EERLEFWMEKYDKDTEMKQNELNALKATKA
1		i				SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK
1						KVKODLLELKSVIKLOAWWRGTMIRREIGGF
						KM
704	2054	Α	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV
						DMSDLSPEEQWRVEHARMHAKHRGHEAMH
						AEMVLILIATLVVAQLLLVQWKQRHPRSYN
						MVTLFQMWVVPLYFTVKLHWWRFLVIWILF
1 1						SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY
-						KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA
						MDFGISLLFYGLYYGVLERDFAEMCADYMA
						STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV
						SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK
1 1						QTCPYCKEKVDLKRMFSNPWERPHVMYGQL
						LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	A	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP
						GNKRPLSALYRLESKEPFLSVGGYVFDYDYY
ł ł						RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA
						VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS
						DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP
						AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ
						IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI
'''				_	,,	S
707	2057	A	5415	6	287	PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT
			1	-	] =	VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP
						EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	A	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS
						TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS
						GCTIFKTVTLTARSTASLLKSVRPRTHQKE
709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG
						SEFTDTILSVHPSDVLDMPVDPNEPTYCLCHQ
				l		VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK
						GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIQPKLSLTLSSSVSRGNVSTPPRHSS
				ŀ		GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY
1 1	i	1		ľ		EESDSDESWTTESAISSEAILSSMCMNGGEEK
		i				PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI
			12			RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV
						SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP
		ļ			ļ	DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL
j		1				NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD
712	2062	_	5400	01	740	LYRDIPELQGF
712	2062	Α	5499	91	749 .	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL
	-	l	1	ļ	ļ	QVRGEALSEEFIWSLLFLAAEQLLEDLRNDSS
L		l				DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KAPELLQGQSEDEQPDASQMHVYSLGMTLY
						WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
713	2063	Α	550 <del>6</del>	22	478	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL GGVECLLLAVMAYDRCVAICKPLHYMVIMN PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR CGHHEVDHFLCEMPALIRMACISTV
714	2064	A	5514	25	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR RLTSSVSCALDEAAAALTRMRAESTANAGQS DK
715	2065	A	5526	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM GRTALFHHSGGSSGYESLRRDSEATGSASSAP DSMSESGAASPGARTRSLKSPKKRATGLQRR RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG SLKEPFEIKVYEIDDVERLQRPRPTTREAPTQG LACVSTRLRLAERRQQRLREVQAKHKHLCEE LAETQGRLMLEPGRWLEQFEVDPELEPESAE YLAALERATAALEQCVNLCKAHVMMVTCFD ISVAASAAIPGPQEVDV
716	2066	A	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD SSSSERTMDLVLEMCNTNSIHWCGISGRQLG KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS VNIREELKLANVVFFPRCLLVQRCGGNCGCG TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2068	A	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS KAEEAAOENA
719	2069	A	5598	ì	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM LATNSTRGLNEDELMAHGQEKDSSSESEDSC PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL FPLAFGLFNIVAAERC
720	2070		5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP LLLRCKLLPSALQKMALGMFFGFTSVIVAGV LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH CPKDFGNINNCRMDLYFFLLAGIQAVTALLF VWIAGRYERASQGPASHSRFSRDRG
721	2071	A	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA KLHETALHHAAKVKNVDLIEMLIEFGGNIYA RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID YLSYN
722	2072	A	5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD VALPALGAVLRGAVRVEKNQELCHLSTIDW GLLQPAPGANHIVGNKLGEECADVCPGVLGA AGEPCAKTTFSGHTDYRCWTSHCQRVCPCP HGMACTARGECCHTECLGGCSQPEDPRACV ACRHLYFQGACLWACPPGTYQYESWRCVTA ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN KAEINPRTNGDRAACQTRTLRFVSNVTEADRI LLRWERYEPLEARDLLSFIVYYKESPFQNATE HVGPDACGTQSWNLLDVELPLSRTQEPGVTL ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS- PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW KPPTQRNGNLTYYLVLWQRLAEDGDLYLND YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH NAITIPISPWKVTSINKSPQRDSGRHRRAAGPL RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY RIDIHACNHAAHTVGCSAATFVFARTMPHRE ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL ILKYEIKYRLGEEATVLCVSRLRYAKFGGV HLALLPPGNYSARVRATSLAGNGSWTDSVAF YILGPEEDAGGLHVLLTATPVGLTLLIVLAA LGFFYGKKRNRTLYASVNPEYFSASDMYVPD EWEVPREQISIIRELGQGSFGMVYEGLARGLE AGEESTPVALKTVNELASPRECIEFLKEASVM KAFKCHHVVRLLGVVSQGQPTLVIMELMTR GDLKSHLRSLRPEAENNPGLPQPALGEMIQM AGEIADGMAYLAANKFVHRDLAARNCMVSQ DFTVKIGDFGMTRDVYETDYYRKGGKGLLP VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP TPRDCSPQNGGPGH
723	2073	A	5672	1	216	LAWIDNILPEKEKKETDKKRKKKGAHEDCD EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR SSSWSYSL
724	2074	A	5704	4235	940	ARGRSRPVWAASWGGRGRPAARRRPRGLA ATMGFELDRFDGDVDPDLKCALCHKVLEDP LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR LSAKELNHVLPLKRLILKLDIKCAYATRGCGR VVKLQQLPEHLERCDFAPARCRHAGCGQVLL RRDVEAHMRDACDARPVGRCQEGCGLPLTH GEQRAGGHCCARALRAHNGALQARLGALHK ALKKEALRAGKREKSLVAQLAAAQLELQMT ALRYQKKFTEYSARLDSLSRCVAAPPGGKGE ETKSLTLVLHRDSGSLGFNIIGGRPSVDNHDG SSSEGIFVSKIVDSGPAAKEGGLQIHDRIIEVN GRDLSRATHDQAVEAFKTAKEPIVVQVLRRT PRTKMFTPPSESQLVDTGTQTDITFEHIMALT KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI GDIHQEMDREELELEEVDLYRMNSQDKLGLT VCYRTDDEDDIGIYISEIDPNSLAKDGRIREG DRIIQINGIEVQNREEAVALLTSEENKNFSLLI ARAELQLDEG WMDDDRNDFLDDLHMDMLE EQHHQAMQFTASVLQQKKHDEDGGTTDTAT ILSNQHEKDSGVGRTDESTRNDESSEQENNG DDATASSNPLAGQRKLTCSQDTLGSGDLPFS NESFISADCTDADYLGIPVDECERFRELLELK CQVKSATPYGLYYPSGPLDAGKSDPESVDKE LELLNEELRSIELECLSIVRAHKMQQLKEQYR ESWMLHNSGFRNYNTSIDVRRHELSDITELPE KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  AAEGISCPSSEGAVGTTEAYGPASKNLLSITE
						DPEVGTPTYSPSLKELDPNQPLESKERRASDG SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR DRLLRERALKIREERSGMTTDDDAVSEMKM GRYWSKEERKQHLVKAKEQRRREEFMMQSR LDCLKEQQAADDRKEMNILELSHKKMMKKR NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF LSVTTV
725	2075	A	5707	3	1770	QISTEVSEAPVANDKPKTLVVKVQKKAADLP DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP WNTDRCFSNYSMVNTTINTSAVVEFWERN MHQMTDGLDKPGQIR WPLAITLAIAWILVYF CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV AASGPGLAFLAYPEAVTQLPISPLWAILFFSM LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR ELFIAAVCIISYLIGLSNITYGGIYVFKLFDYYS ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ MTPLTMGNYVFPKWGQGVGWLMALSSMVL IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI
726	2076	A	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII YCNDGFCEMTGFSRPDVMQKPCTCD
727	2077	A	5716	3		HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL AWFEKMTCYLQLLFNICLPDVSEE
728	2078	A	5737	1899		IQASRASPYPRVKVDFALSCHEDLLAPISEPIE WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN EEVLADVRTIVNQISYTPQDPRDLCGRILTTC YMASKNSSQETCTRARELAQQIGSHHISLNID PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE LEPLADGQVSQTDEEDMGMTYAELSVYGKL RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE DNRFDLRPFLYNTSWPWQFRCIENQVLQLER AEPQSLDGVD
729	2079	A	5741		5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFQTPPRCPVPLL LLLLGAARAGALEIQRRFPSPTPTNNFALDG AAGTVYLAAVNRLYQLSGANLSLEAEAAVG PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide scq- ucance	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHP SDPPPGAQSYAYLALNSEARAGDKESQARSL LARICLPHGAGGDAKKLTESYIQLGLQCAGG AGRGDLYSRLVSVFPARERLFAVFERPQGSPA ARAAPAALCAFRFADVRAAIRAARTACFVEP APDVVAVLDSVVQGTGPACERKLNIQLQPEQ LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV AVASVNNYTAVFLGTVNGRLLKINLNESMQ VVSRRVVTVAYGEPVHHVMQFDPADSGYLY LMTSHQMARVKVAACNVHSTCGDCVGAAD AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ IAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA
						AYCGWCALETRCTLQQDCTNSSQQHFWTSA SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ IAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN
						RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI HGNDLHVGSELQVLVNDTDPCTELMRTDTSI ACTMPEGALPAPVPVCVRFERRGCVHGNLTF WYMQNPVITAISPRRSPVSGGRTITVAGERFH MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP GALSNASAPVDFFINGRAYADEVAVAEELLD
						PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV LLLLSVVALFVFCTKSRRAERYWQKTLLQME EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSQT LNSQGSSQAQETHPLLGEWKIPESCRPNMEE
						GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD RCSLASLLTIALHGKLEYYTSIMKELLVDLID ASAAKNPKLMLRRTESVVEKMLTNWMSICM YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG KARYTLNEEWLLRENIEAKPRNLNVSFQGCG MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY SQWPRAEDVDLEWFASSTQSYILRDLDDTSV VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH
					,	RQSHRKKVLPEIYLTRLLSTKGTLQKFLDDLF KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD PDTLHIWKTNSLPLRFWVNILKNPOFVFDIDK TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEQE MNAHLAEESRKYQNEFNTNVAMAEIYKYAK RYRPQIMAALEANPTARRTQLQHKFEQVVAL MEDNIYECYSEA
730	2080	A	5744	3	292	QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ VANGKGNQRNMGSPQPSLLAFERNLELQIMG LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT LKD
731	2081	A	5747	1	382	FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	1=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	- [	ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-		Ì		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion, \=possible
	<del>-</del>			sequence		nucleotide insertion
1	1	i				LSFSLRSSRVSGRHWKNFALVPLLREASARD
	1	ļ			1	RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS GEK
732	2082	A	5753	198	3	AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
/32	. 2002	^	3/33	170	] 3	PGKLQAQLPCPSPVRFTSARIPPASRPOTKS
733	2083	A	5754	2	2223	AAGPPGLEAEGRAPESAGPGPGGDAAETPGL
( ,33	2005	l ^ _	3734	-	2223	PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL
	Ì	1	i			SVANCLGAQTVQAPAEPAAGKAEQGETSGR
		1	ļ			EAPEAPAVGREDASAEDSCAEAGASGAADG
	1	l	}		ļ	ATAPKTEEEEEEEETAEVGRGAEAEAGDLEO
	1					LNRTSTSTKSAKSGSEASASASKDALQAMILS
						LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN
		ŀ				LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK
	1					GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM
						DFSSMELDEALRKFQAHIRVQGEAQKVERLIE
						AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL
						NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG
				'		ADIPRELVVGIYERIQQKELKSNEDHVTYVTK
			1			VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV
						NKLQKQAAHQREVFLFNDLLVILKLCPKKKS
						SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV
1						TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA
	i i		1			QGDPQSKQGSPTAKREAALRERPAESTVEVSI
						HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR
	1					QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC
	1					LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG
ì	1		İ .		·	SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH
						GHRYSSGSRSLV
734	2084	Α	5788	8	362	SSVMGDLVGQGLEEQIVARDENSWLIDGGTP
						IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR
1	1					KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT
						RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL
i						PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST
1		ļ ,				AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG
						NLVVCLMVYQKAAMRSAINILLASLAFADM
						LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF
				-		FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR
	]					PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY
	•					SFMGILNTLRHNALRIHSYPEGICLSQASKLGL
			) [			MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC
						WAPFTTYSLVATPSKHFYYQHNFFEISTWLL
						WLCYLKSALNPLIYYWRIKKFHDACLDMMP
						KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKOFSRS
				l		DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP
						LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR
						RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH
1				İ		GLARRLTTARRPPASSEQAQQELFNELKPAV
		l		ļ		DGANFIVNHMRDQNNYNEEKDSWNRVART
i	1			ł	ı	VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP
						QPFPGDPYSYNVQDKRFI
738			7004	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCG
1	2088	A	5881	7	1100	
	2088	A	5881	1	1100	LLDSSKLCDYENRFNTSKGGELPDRPAGVGV
	2088	A	5881	1	1100	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ĺ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł		ļ	peptide		/=possible nucleotide deletion, \=possible
	L	1		sequence	1	nucleotide insertion
						WCSQGADCITPGLYAMVGAAACLGGVTRMT
l	[		l			VSLVVIMFELTGGLEYTVPLMAAAMTSKWVA
ŀ					ļ	DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL
					·	AMDVMKPRRNDPLLTVLTQDSMTVEDVETII
		1				SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE
		İ				NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK
1		[	1 1		ĺ	LRNILDLSPFTVTDLTPMEIVVDIFRKLGLRQC
		1				LVTHNGRLLGIITKKDVLKHIAQMANODPDSI
	·					LFN
739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
		}				DQALQELRKVARINGHKEAKNLTIEVLMSSV
		1				KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV
		]				VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA
						VDFLGRATTALLLSFLGRRTIQAGSQAMAGL
			(	1		AILANMLVPQDLQTLRVVFAVLGKGCFGISL
						TCLTIYKAELFPTPVRMTADGILHTVGRLGA
						MMGPLILMSRQALPLLPPLLYGVISIASSLVVL
						FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE
			4			AFTVESTSLLEIVALHGAL
740	2090	Α	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE
						NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG
1						KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM
						GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV
						CSEILLSRMCANKRTM
741	2091	Α	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV
						VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER
						QRELKEKIREERRNKLAAEMGEDGEKEFQEE
1						EEEKEEEEEEEPLPEIFIPSTPSPILCGFYSEPG
						KFWV
742	2092	Α	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV
1 1						TQMGNDKSIKCEQNLGHDTMYWYKQDSKK
						FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL
						NLHINSLELGDSAVYFCASSQDTALQSHCIPV
] ;		Ì				HKPPGSARKLQGSVCTCTQGSSLHSLMASDG
						VPVC
743	2093	Α	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER
]	1			[		RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
}				0.0		GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
	ļ	J		,	j	ADRARRERFIMNEKWDTNSSENWHPIWNVN
						DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF
		.	[			LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI
	]					LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM
	1	ĺ	- 1	ſ		CKISGLVQGISVAASVFTLVAIAVDRFQCVVY
	j					PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH
	l	l				VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ
	ļ	·		1		EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
	j	J				RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI
		- 1			ı	VALLFILSWLPLWTLMMLSDYADLSPNELOII
				İ		NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG
	1		ļ	l		FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN
	1	ł	1	1	ł	TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
		ــــــــــــــــــــــــــــــــــــــ			!	LVMEELKETTNSSEI
744	2094	A	5966	149	327	SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
						LYDYQGGRLGVARGAWYMEAPDIRQGDM
745	2095	A	5970	413	856	GAPHTDWAWAPTPMSGLGSGRGRQGTLASS
	1	*	i	· .	1	PLSLPLLLAGVTGILATELFDQMARPAACMV
-[	j				l	CGALMWIMLILVGLGFPFIMEALSHFLYVPFL
		İ		ľ		GVCVCGAIYTGLFLPETKGKTFQEISKELHRL
			1			NFPRRAQGPTWRSLEVIQSTEL
					·	

	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
- 1	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
- 1	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
- 1	cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
	seq-	neuce	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
1	uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
-1			ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			İ		peptide		/=possible nucleotide deletion, \=possible
L			Ĺ		sequence		nucleotide insertion
١	746	2096	Α	5971	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
			l		<b>.</b>		RCARHGACQRSCLASQDPYCGWHSSRGCVDI
1					ŀ	}	RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
			l				SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
-							AAAFALGASVSGLLVSCACRRAHRRRGKDIE
					J	j .	TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
	j		İ				VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
			l				LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
							GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
1							EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
	Į			ļ	,		APALLGGPSPRPHECASPLRLDVPPEGRCASA
			l				PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
			l				LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
			1		İ		RALKRVDVEKPQLSLKPPLVGPSSRQAVPNG
Ĺ			l	L	<u> </u>		GRFNF
Γ	747	2097	Α	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGQVTI
							LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
			l	1			RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
			ŀ				DRVQALSYAQHTRQLISCGGDGGIVVWNMD
1	- 1		· .	l			VERQETPEWLDSDSCQKCDQPFFWNFKQMW
ŀ							DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
	l		İ				MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
4	i			1			HNTVHVHFDATRGWLLTSGTDKVIKLWDMT
L							PVVS
Γ	748	2098	Α	6001	2 .	747	AMVFGGVVPYVPQYRDIRRTQNADGFSTYV
							CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
							TMLLMLKLCTEVRVANELNARRRSFTAADS
1	!			l i			KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
1	}						QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
							LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
Į							WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
1				·			DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
Г	749	2099	A	6002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
1							RARLADDLNEKIAQRPGPMELVEKNILPVDSS
							VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
1	1	-					DQPASQESQGSAASPSEPKVSESPSPVTTNTP
							AQFASVSPTVPEFLKTPPTAD
1	750	2100	A	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
							WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
				[	ı		QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
	į	ļ					WDLKRDVAKKLEKLLKRTORAIAELIRERLK
1	1						GQEDSLDSAVDAATEHKTC
$\vdash$	751	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
							SHPDKLKRMSKSVPAFLQDESDDRETDTASE
		1			.		SSYOLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
	l						VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
	l			-			HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
	ł	ŀ		/ I		ľ	DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
ļ							QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
	1				ļ [		TWDWDNKQNKQLRWYPLKRKTAPVALEAE
	1	}				ļ	NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
1					<b> </b>		KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
	ŀ						KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
1		ļ				İ	
1	*	- 1					VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT EVDWMDSTSEEVALWEKMVNSPNTWIEATL
1				. 1			
	j	j		) }	1	1	DIDIATIIAKICK
-	752	2102		6020	100	1282	PLRMLLIAKISK VETESPEELISVERI CLI LGVTCSOSMAESELI
-	752	2102	A	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL
-	752	2102	A	6028	108	1283	

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV
						SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK KNLKEKA
753	2103	A	6043		1470	DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEFA FGSKENTLNIKKALLSGYFMQIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	A	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRPTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGQGCPGVAPEVTEGAKGLEDTEE PEEEEBGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106		6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	A	6063	54	419 .	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

SEQ ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		[	1	peptide		/=possible nucleotide deletion, \=possible
		L	<u> </u>	sequence	L	nucleotide insertion
						QILTMLLRSLQQPSASWPRDCSSSCSW
758	2108	A	6066	125	438	IGISCPATIFVPMFSHSLIGIGEEYQLPYYNMV
			1	1		PSDPSYEDMREVVCVKRLRPIVSNRWNSDEC
						LRAVLKLMSECWAHNPASRLTALRIKKTLAK
						MVESQDVKI
759	2109	Α	6072	3	650	PGRRFRPAALEERAMEKLREKVPFQNRGKGT
						LSSIIPNNSDTRKATETTSLSSKPEYVNPDFRW
, ,						SKDPSSKSGNLLETSEVGWTSNPEELDPIRLA
					ļ	LLGKSGLSCQVGSATSHPVSCQEPIDEDQRISP
						KDKSTAGREFSGQVSHQTTSENQCTPIPSSTV
						HSSVADMQNMPAAVHALLTQPSLSAAPFAQ
760	2110	Α	6077	3 ·	730	RYLGTLPSTGSTTLPQCHAGNATVW
/00	2110	A	00//	3 .	/30	PLRLTLMEEVLLLGLKDREGYTSFWNDCISSG
						LRGCMLIELPLRGRLQLEACGMRRKSLLTRK VICKSDAPTGDVLLDEALKHVKETQPPETVQ
						NWIELLSGETWNPLKLHYQLRNVRERLAKNL
						VEKGVLTTEKQNFLLFDMTTHPLTNNNIKOR
						LIKKVQEAVLDKWVNDPHRMDRRLLALIYL
						AHASDVLENAFAPLLDEQYDLATKRVROLLD
				1		LDPEVECLKANTNEVLWAVVAAFTK
761	2111	A	6078	833	390	IVSFHLSGFKKFVRPFSFLSVHGLQVDEYHSV
				300		HQKLSADMADHSNLIRSLLVGAEDARLMRD
] .						MKTMKSRYMELYDLNRDLLNGYKIRWNNH
						TELLGNLKAVNQAIQRAGRLRVGKPKNQVIT
						ACRDAIRSNNINTLFKIMRVGTASS
762	2112	Α	6079	2	2686	KKAITCGEKEKQDLIKSLAMLKDGFRTDRGS
						HSDLWSSSSSLESSSFPLPKQYLDVSSQTDISG
					1	SFGINSNNQLAEKVRLRLRYEEAKRRIANLKI
						QLAKLDSEAWPGVLDSERDRLILINEKEELLK
						EMRFISPRKWTQGEVEQLEMARKRLEKDLQ
						AARDTQSKALTERLKLNSKRNQLVRELEEAT
				]	·	RQVATLHSQLKSLSSSMQSLSSGSSPGSLTSSR
						GSLVASSLDSSTSASFTDLYYDPFEQLDSELQ
]						SKVEFLLLEGATGFRPSGCITTIHEDEVAKTQ
[						KAEGGGRLQALRSLSGTPKSMTSLSPRSSLSS
						PSPPCSPLMADPLLAGDAFLNSLEFEDPELSA
.	ļ					TLCELSLGNSAQERYRLEEPGTEGKQLGQAV
		.				NTAQGCGLKVACVSAAVSDESVAGDSGVYE
}			· l			ASVQRLGASEAAAFDSDESEAVGATRIQIALK
		- 1				YDEKNKQFAILIIQLSNLSALLQQQDQKVNIR
		- 1				VAVLPCSESTTCLFRTRPLDASDTLVFNEVFW VSMSYPALHQKTLRVDVCTTDRSHLEECLGG
		- 1				AQISLAEVCRSGERSTRWYNLLSYKYLKKOS
.	ſ	- 1	1			RELKPVGVMAPASGPASTDAVSALLEQTAVE
		i				LEKRQEGRSSTQTLEDSWRYEETSENEAVAE
						EEEEEVEEEEGEEDVFTEKASPDMDGYPALK
						VDKETNTETPAPSPTVVRPKDRRVGTPSOGPF
			- 1			LRGSTURSKTFSPGPQSQYVCRLNRSDSDSST
						LSKKPPFVRNSLERRSVRMKRPSPPPQPSSVK
				1		SLRSERLIRTSLDLELDLQATRTWHSQLTQEIS
}		ļ		1	,	VLKELKEQLEQAKSHGEKELPQWLREDERFR
	. [	[				LLLRMLEKRMDRAEHMGELQTDKMMRAAA
		.		İ	1	KDVHRLRGQSCKEPPEVQSFREKMAFFTRPR
					•	MNIPALSADDV
763	2113	A	6082	3	1558	PHPIRFSKLCVSFNNQEYNQFCVIEEASKANE
						VLENLTQGKMCLVPGKTRKLLFKFVAKTED
			1			VGKKIEITSVDLALGNETGRCVVLNWQGGGG
		l-		ļ	1.5	DAASSQEALQAARSFKRRPKLPDNEVHWGSII
			ł	}		IQASTMIISRVPNISVHLLHEPPALTNEMYCLV
			J			-4

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK
						THVTLHGTELCDESYPALLTDIPVGDLHPGEQ LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF EHLERVYADIPFLLMTDLLSASPWALTIVSSE LHLAPSMTTVDQLESQVDNVILQTGESASECF CLQCPSLGNIEGGVATGHYIISWKRTSAMENI PIITTVITLPHVIVENIPLHVNADLPSFGRVRES LPVKYHLQNKTDLVQDVEISVEPSDAFMFSG LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM DDTSIAAA
764	2114	A	6093	1	1422	AAADLANSNAGAAVGRKAGPRSPPSAPAPAP PPPAPAPPTLGNNHQESPGWRCCRPTLRERN ALMFNNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA AFLILLKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA
765	2115	A	6099	1		SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK
766	2116	A	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEBLYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	Α .	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A.	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSS SSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY KHEDLQTDESSMDDRHPRRQLCGGNQAATE

(TONO TO	000 10	1. 3. C. A	Laro	T 8 10 10 10 10 10 10 10 10 10 10 10 10 10	1 8 1 1 2 1 2	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-	Į	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	}		914	ng to first	acid residue	Q=Ghttamine, R=Arginine, S=Serine,
700		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ł	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1		peptide		/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
						RIILFGRELQALSEQLGREYGKNLAHTEMLQD
l	ł	l	1	1		AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
1	ŀ	1	1		1	NSAILESQNLPKQPPLMLALGQASECLRLMA
						RAGLGSCSFARVDDYLH
770	2120	A	6125	2	570	YFGLNLHVQHLGNNVFLLQTLFGAVILLANC
1	1	ł			ł	VAPWALKYMNRRASQMLLMFLLAICLLAIIF
	1	Į				VPQEMQMLREVLATLGLGASALANTLAFAH
						GNEVIPTURARAMGINATFANIAGALAPLMM
				ĺ	ł	ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
1	1	ľ	ì		ĺ	PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
771	2121	A	6126	909	353	
1 ′′′	2121	I A	6126	909	333	RSFVLDTASAICNYNAHYKNHPKYWCRGYF
1	ļ				1	RDYCNILAFSPNSTNHVALRDTGNQLIVTMSC
1	ĺ	(	i		ĺ	LTKEDTGWYWCGIQRDFARDDMDFTELIVT
				·		DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
}	1	)			ļ	RKADRSRTSILIICILITGLGIISVISHLTKRRRS
			J	l		QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	Α	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
						TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
1	ļ	1	i	]	}	WYLLRKHWIANNLFGLAFSLNGVELLHLNN
	1	1				VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
ľ	i					FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
			1		İ	VIPGIFIALLLRFDISLKKNTHTYFYTSFAAYIF
ļ	ì		1			GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV
		1	1			ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
	1		1			EASASKGLEKKEK
773	2123	A	6161	3	1088	COPMLYTRKNHPKLLLRRTESVAEKMLTNW
1 '''	2120	1 ^	0.01	, ,	1000	FTFLLYKFLKESAGEPLFMLYCAIKHOMEKG
1	1					PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
ł		1	1			
						NPENENAPEVPVKGLDCDTGTQAKEKLLDA
ĺ	ł	ł	1 .	•		AYKGVPYSQRPKAADMDLEWRQGRMARIL
	İ					QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
						ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
	٠ .					SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
1.	ł	l	ł			LDQREGDRGSKMVSEIYLTRLLATKGTLQKF
1			1			VDDLFETIFSTAHRGSALPLAIKYMFDFLDEQ
						ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
1	1		1			NPQFVFDIHKNSITDACLSVV
774	2124	Α	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
		1				SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
		ł				PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
	1	l				GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
		Į.				DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
		l			ĺ	NHTMVYDGFGPADLRQACAELSLWDHGALA
		l				NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
						QLWQALLEQPCEWVDGLLPLRTNLAPRT
775	2125	A	6191	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
113	2123	^	1610		374	
		1		-		YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
		1		[ .		DMK\KSPEIISRRMTFAL*CYSLTFVRFAHYVQ
[		ĺ	[	·	ľ	\PWNWLMLGCHTAVDFDQLISSMPCISHGMT
·						ASASAL
776	2126	A	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
		1				RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
[		[	( [			QAALRGRIQLREQLPRYLMGETQLCTTRCVL
						KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
	\				ı	VSGSECLLSMVEEGNPHHYFVATODONLSVK
[						VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
			[		Ì	VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
j :						RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
[ ·			<b>!</b>			KKRKRKRIRNRSNPKVLSEKONAEGE
L		L	<u> </u>			MANAGER A DIENVINEUE

Concern	I CEA VE	I Nach	Loro	D., 31.4. 1	T N 21 2	I A mino cold comment (A. Alimir C. C.
SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
NO: of	NO: of	поа	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,   F=Phenylalanine, G=Glvcine, H=Histidine.
eotide	,	i	_	location		
seq-	seq-	]	USSN 09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
nence	uance	Į	914	ag to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciice		i	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide		/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
1		1	0200	1000		FQPFGRPRGGGHLRSGVLGQPGOHGETP/SFF
		l				YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ
						RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	Α	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
	1	ŀ	į į			AVSNADILLDLKIN*LDHILEEKMPLIEVKVVP
		1				PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
						SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
1		]				LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
			L			NSFRYRR
779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
		1				YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
		1		00		FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
181	1					MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
						QSQPMY
780	2130	Α	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
	1	[				TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
1						HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
						QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
	1	l				AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
	1	]				S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
		1	1			FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
1	1					LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
		1	]			PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
	1	i i				PV PV
781	2131	A	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
/01	2.31	^	0217	332	210	LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
	1	ŀ			:	VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
1	ł	1			İ	QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
	1		[ 1			LEFLGDLKGCSELKNFQELITQSALVHPKADV
						WWYCGRPLLGTLPSN
782	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
				-		EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
						DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
						KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
						DDKKRVKAKKKKKKKKKKKKKKKKKKKKKK
						ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
1	<b>[</b>					IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA
1						EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
1				;		SGSRHRRMEAVRLRKENQIYSADEKRALASF
<u></u>		ليسيا				NQEERRKRESKILASFREMVHKKTKGKDDK
783	2133	A	6305	201	1032	WDDYPQGALRREAAEGLHFLGPPGRVRGQ
				l	l	LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
1				l	ļ	AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
	<b> </b>			l	į	AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
1	]			l		SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
1				l		TTVVFWPAKLQASSRVVMFRFEFWDCGESA
	<b>!</b>		' I	ľ		LKKFDHMLLACMENTDAFLFLFSFTDRASFE
1				l		DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
70.4	10154	ĻI	(222			DVPERDLTAFRQAWELPLLRVKSVPGRRLG
784	2134	A	6308	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP
ł	į į				ł	GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
	]					PRKPEGAQARTAQSGALRDVSEELSRQLEDIL
						STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
		1				TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR
		<b> </b>	1			QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
1		ļ	]			QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA
L						NAME TO SELECT A CENTURE OF SELECTION AND A SELECTION AND A SELECTION AS A SELECT

COPO ID	L CEO TO	1 1/4	Loro	T D 1 4 1	18 0:1-1	
SEQ ID	SEQ ID	Met hod	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	поц	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	peptide		in		location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1 .	1	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ŀ	1		1	peptide		/=possible nucleotide deletion, \-possible
<u> </u>	ļ	<b> </b>	<u> </u>	sequence		nucleotide insertion
1	1		İ			RSKLESLCRELQRHNRSLKEEGVQRAREEEE
	İ	ļ		1	•	KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR
1	1	1	ł	ļ	ŀ	QENMELAERLKKLIEQYELREEHIDKVFKHK
1		1	1			DLQQQLVDAKLQQAQEMLKEAEERHQREKD
1			1			FLLKEAVESQRMCELMKQQETHLKQQLALY
į	Į.	ł				TEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI
1	1 .	ĺ	l I		[	KKLEKETTMYRSRWESSNKALLEMAEEKTV
	1	1				RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
						GQRWGSHRTSAVRIFS
785	2135	Α	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT
İ		1	Į.			PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT
1	1					KLPWSWGMRPMKIFFSEEYRSISTRISHDAL*
1						EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT
1		1	'			RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S
1		1	1		l ,	QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD
		1				TTPCOKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA
1,00	1 2130	1	0320	331	133	REHGQCADVDECSLAEKTCVRKNENCYNTP
		1	1			
1	1	l	1			GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK
	1		1			PDTAALPRRPVMCRTYPLNYSEGCPVENVAL
787	2137	<del>                                     </del>	6220	1600	000	RMPSPAVDSGGERLPAL
/8/	2137	Α	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG
1	-	i	1 1			SGILGLAYVMANTGVFGFSFLLLTVALLASYS
1						VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
1						PLIEFLQSL*NSL\*AVTSYEDLGLFAFGLPGKL
1		}				VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT
i		1				GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG
í	ĺ	ĺ				FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL
		j				NYVEKGFQISNVTDDCKPKLFHFSKESAYALP
	1					TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN
1			]			TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE
		ļ				VTCHRIKDKVESELLKG***IP*SHDVVVMT\V
		1				KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP
	1	1				FSWIRHFLITLALNIIIVLLATYVPDIRNVFGVV
1	i	ĺ	i i			GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV
			!			GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351	1	6622	PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM
		1				SASFVPNGASLEDCHCNLFCLADLTGIKWKK
			<u> </u>		·	YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
			1 1	i		LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF
						TMTYQKKKMECGRMDFPMNAVLCFSKAVH
					l	NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN
1					ļ	KSEHLSCSFTFFLHGDSNVCTSVEINQHOPVY
					ļ	LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ
1				ļ		AFKMSDSATKKLIGEWKOFYPISCCLKEMSE
						EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC
]					ļ	FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS
				İ		TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW
						VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV
1						DRVWQECNMNRAQNKRKYSASSGGLCEEAT
1 1	l				ł	AAKVASWDFVEATQRTNCSCLRHKNLKSRN
						AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK
'						PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL
			•		. 1	V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE
						MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS
			! I	ł	1	QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ
	0.			-	1	YQEAVEPTVYVGTAVNLEEDEANIAWKYYK
				1	1	FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV
]				1	ļ	TSVTELMVOCKKPLKVSDELVOOYOIKNOCL
				1	1	SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
789	2139	A	6359		2002	RQNSEREAGKKHKVEDGTSSVTVLSHEEDA MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA VSYTDLDNLFNSDEDELTPGSKRSANGSDDK ASCKESKTGNLDPLSCISTADLHKMYPTPPSL EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY KPENCQILVGCSMFAPLKTLPSQYLPLIKLPEE CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI LPSFSTPRFPTPRTPRTPRTPRTGAGGPASAQGS VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP EAHSLYVNLILSESVMNLFKDCNSDSCCICVC NMNIKGADVGVYJPDPTQEAQYRCTCGFSAV MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE KRFEALRATSAEHVNGGLKESEKLSDDLILLL QDQCTNLFSPFGAADQDPFPKSGVISNWVRV EERDCCNDCYLALEHGRQFMDNMSGGKVDE ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV LSPFALPYWERLMLEPYGSQRDIAYVVLCPE NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS SLLSQPNLVAPTSQSALTPPQMTNTGNANTPS ATLASAASSTMTVTSGVAISTSVATANSTLTT ASTSSSSSSNLNSGVSSNKLPSFPPFGSMNSNA AGSMSTQANTVQSGQLGGQQTSALQTAGISG ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL GLLRCFLEMVQTLPPHIKSTVSVQIPCCYLLQ PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS TNVKTLTGFGFGLAMETALRSPDRPECIRLYA PPFILAPVKDKQTELGETFGEAGQKYNVLFV GYCLSHDQRWILASCTDLYGELLETCIINIDVP NRARKKSSARKFGLQKLWEWCLGLVQMSS LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ SLSKRLKDMCRMCGISAADSPSILSACLVAM EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL NTPQDTSCTHILVFPTSASVQVASATYTTENL DLAFNPNNDGADGMGIFDLLDTGDDLDPDII NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP LPDWFWSACPQAQYQCPLFLKASLHLHVPSV QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI MNML TGTLTEDGLDVMGVVPLKGOAFLPLVPEPRR
	213 <b>9</b> .	A -	<b>633</b> Y	1		TGILTEDGLDVMGVVPLKGQAFLPLVPEPRR LPVGPLLRALATCHALSRLQDTPVGDPMDLK MVESTGWVLEEEPAADSAFGTQVLAVMRPP LWEPQLQAMEEPPVPVSVLHRPFSSALQRM SVVVAWPGATQPEAYVKGSPELVAGLCNPET VPTDFAQMLQSYTAAGYRVVALASKPLPSVP SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA RGCGMVAPQEHLIIVHATHPERGQPASLEFLP MESPTAVNGVKDPPQAASYTVEPDPRSRHLA LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP EQKTELVCELQKLQYCVGMCGDGANDCGAL KAADVGISLSQAEASVVSPFTSSMASIECVPM

SAU ID Not of moch coilide services of the peptide of the peptide of the peptide of the peptide of the peptide of peptide of peptide of peptide of peptide of peptide of peptide of peptide of the peptide of p	C 070 00	T 000 00		1 000		( <del>-</del>	1
in uncleotide sequence USSN location corresponding to the properties where the properties of the prope	SEQ ID	SEQID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence by 14 by 15	4		род		beginning		
Sequence   1914   1							
Page   Page		, -			1		
mino acid residue of peptide residue of peptide residue of peptide sequence physics sequence public sequence propriet sequence provide sequence provide sequence provide sequence provide sequence provide sequence provide sequence provided sequence	•	uence	1				
residue of popition   popition	uence	1	ĺ	914			
peptide sequence    Possible nucleotide deletion, impossible nucleotide insertion   nucleotide insertion   nucleotide insertion   viriant.   VIREGRESI.DTSFSVFKYMALYSI.TOFISULI.   VIREGRESI.DTSFSVFKYMALYSI.TOFISULI.   VIREGRESI.DTSFSVFKYMALYSI.TOFISULI.   VIREGRESI.DTSFSVFVRJARTAVABADILPNY   ENTVYESI.SSFOYLILAAAVSKGAPFRUPLITN   NVPFILASALP.SSVLVVI.SSPCILHGELAE.PNW   PENTVYESI.SSFOYLILAAAVSKGAPFRUPLITN   NVPFILASALP.SSVLVVI.SSPCILHGELAE.PNW   PENTVYESI.SSFOYLILAAAVSKGAPFRUPLITN   NVPFILASALP.SSVLVVI.SSPCILHGELAE.PNW   PENTAGRES.PSFASKUVVI.SSPCILHGELAE.PNW   PENTAGRES.PSFASKUVVI.SSPCILHGELAE.PNW   PENTAGRES.PSFASKUVVI.SSPCILHGELAE.PNW   PENTAGRES.PSFASKUVVI.SSPCILHGELAE.PNW   PENTAGRES.PSFASKUVVI.SSPCILHGELAE.PNW.DATA   NVEELLEQADVI.YSSGETEKI.VOI.LTOYKSES   DAEILWRI.ASAS.PNW.GLSINTSESKI.LVV   EALEY AKRA/LSKNESSFASKUV.PNW.GLS.PSFASKUV.PNW.GLS.PSFASKUV.PNW.GLS.PSFASKUV.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKUV.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNW.GTASALPYPYORGRANACI.QIP   PEPPTYSK.ALGVIPRAGRES.PSFASKU.PNG.PSFASKU.PNG.GTASALPYPASKU.PNG.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANACI.QIP   PEPPTYSK.ALGVIPRAGRANA	ľ	1	l			1 1 1	
Sequence   nucleotide insertion	1		ļ	1		sequence	
VIREGRESLIDTSFSVFKYMALYSLTOFFSVILL   YTINTLOGLOPHALDUTITYAVALMSRTOP   ALVUGRVRPPGALLSVPVLSSLLLQMVLVTG   VQLOGYFLTLAQPWEVPLRITYAVADPULPRY   VQLOGYFLTLAQPWEVPLRITYAVADPULPRY   ENTVYFSLSSFOYLILAAAVSKGAPFRUPLTN   NVPFLLASAL*SSVLVVLVSSCLHGPLAE   NITDTGFKLLLVGLVTINFVGGLHAGERARP   VPFLPAPPPAQAGSKKRFKQLERELAEQPW   PPLPAGPLR   NTDTGFKLLLVGLVTINFVGGLHAGERARP   PPLPAGPLR   SSAGSARKLQVMALAARLWRLLFPERGAAP   GSRLPAGTSGSRGHIGPCFFRGFEVAGNPGT   KRIGLLSALS*LSTETTVOVISGAVAVHATA   KVEELLEQADVLYSSGETEKLVQLLTOYKESE   DAELLWRLARASRDVAQLSKTSEEKKLLVY   GALEYARRA/LØKPGSSTASKWVALGAT   VGLELPAGAT	1		1	1			
YTINTNLGDLQPLAIDLVITTTVAVLMSRTGP		<b></b>		<u> </u>	sequence		
ALVLGRVRPPGALLSVPVLSSLILQMYLYTIG VQLGGYFLTLAQPWEYDLRTVAADNLRNY ENTVYFSLSSTQYLILAAAVSKGAPFRUPLTN NVPFLASAL*SSULVLAAAVSKGAPFRUPLTN NVPFLASAL*SSULVLAAAVSKGAPFRUPLTN NVPFLASAL*SSULVLAAAVSKGAPFRUPLTN NVPFLASAL*SSULVLAAAVSKGAPFRUPLTN NVPFLASAL*SSULVJLSSCHEGELAGQPW PPLPAGPLR  790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLLPFRKGAAP SSAGSARKLQVMALAARLWRLLPFRKGAAP SSAGSARKLQVMALAARLWRLLPFRKGAAP SSAGSARKLQVMALAARLWRLLPFRKGAAP SSAGSARKLQVMALAARLWRLLPFRKGAAP SKELEQADMYLSSESFASHKWYACLSDV GDYEGKAKIANAYIKSHFEKAELNPKDATS HILMGINCVTTEAMPWQRALA*NACLQIPP *PPYEKALGYPHRAEQVDPNTYSKNLLLLG GDYEGKAKIANAYIKSHFEKAELNPKDATS HILMGINCVTTEAMPWQRALA*NACLQIP *PPYEKALGYPHRAEQVDPNTYSKNLLLLG GDYEGKAKIANAYIKSHFEKAELNPKDATS HILMGINCVTTEAMPWQRALA*NACLQIP *PPYEKALGYPHRAEQVDPNTYSKNLLLLG KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KTYLKLHNKKLAAPHWGARDYPHYSKNLLLIQ KOQCAGGALALARVYSSISHAN NETSGTYANGSVIDLEKSEGYYNNVSGOPSP DOSGLUDMGINGKQROGPHYDLTSVPHATYSS DOSGLUDMGINGKQROGPHYDLTSVPHATYS DOSGLUDMGINGKQROGPHYDLTSVPHATYS DOSGLUDMGINGKQROGPHYDLTSVPHATYS POSGLUDMGINGKQROGPHYDLTSVPHATYS PNNGQLAFGITMTEIDRIAQNIKSHLETCQY TIMEELHQLAWQTHTYEELGAQNIKSHLETCQY TIMEELHQLAWQTHTYEELGAQNIKSHLCONDQ ILLIKSGCLEVVLVRNCRAFPPLNNTVLFEG KYGGMQMKALGSDLVNGADPSKANLCSL QCACIOTHAIQ VVAFCARMTGMELCONDQ ILLIKSGCLEVVLVRNCRAFPPLNNTVLFEG KYGGMQMKALGSDLVNGADPSKANLCSL QCACIOTHAIQ VVAFCARMTGMELCONDQ ILLIKSGCLEVVLVRNCRAFPLNNTVLFEG KYGGMQMKALGSDLVNGADPSKANLCSL QCACIOTHAIQ VVAFCARMTGMELCONDQ ILLIKSGCLEVVLVRNCRAFPLNNTVLFEG KYGGMQMKALGSDLVNGADPSKANLCSL QCACIOTHAIQ VVAFCARMTGMALCONDQ ILLIKSGCLEVVLVRNCRAFPLNNTVLFEG KYGGMQMKALGNGANLCSNATCH GCACHAPATACH SAGGVALCASSTGNICKSCHANLCSNATCH TOTAL CHARACTACH	1	l	l	ļ			
	1	ŀ	1			i	1
PATTVYFSLSSFQYLLIAANYSKGAFFRRPLAN   NITDTGFKLILVGLVTLNFVGGLHAGERARP   VPPRLPAPPAGASKKRFKQLERELAEQPW   PPLPAPPAGAGSKKRFKQLERELAEQPW   PPLPAGELR   SSAGSARKLQVMALAABLWRLLPFERGAAP   SSAGSARKLQVMALAABLWRLLPFERGAAP   GSRLPAGTSGSRGHCGPCRFRGFEVMGNPGT   FRRGLILSALSYLGFETYQVISQAAVVHATA   KVEELEQADYLYGSGETEKLYQULTQYKESE   DAELLWRLARASKDVAQLSRTSEEEKKLUV   EALEW AKRAJLEMENSSYASHKWYALCLSDV   GDYEGIKAKIANAYIIKEHFEKAIELNFKDATS   HIMMINCYTFAEMPSYAGRIKWYALCLSDV   GGYEGIKAKIANAYIIKEHFEKAIELNFKDATS   HIMMINCYTFAEMPWYQRRIAWNACLQLPP   *PPPYEKALGYFHRAEQVDPNFYSKNLLLIQ   KYTIKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAAQLITSTSEKN   KQIQTEAAQLITSTSKA KQIQTEAAQLITSTSKA KQIQTEAAQLITSTSKA KQIQTEAAQLITSTSKA KQIQTEAAQLITSTSKA KQIQ	1						
NVFFLIASAL-SSVLVVILVSGLHGGHAGERAR   NITDTGKLILVGLVTLNFVGGHAGERAR   NITDTGKLILVGLVTLNFVGGHAGERAR   NITDTGKLILVGLVTLNFVGGHAGERAR   VPPRI PAPPAQAG\SKKRFKQLERELAEQPW   PPLPAGPLR   SSAGSARKLQVMALAARLWRLLPFRGAAP   GSRLPAGTSGSRGHGCPCRRGGEVMGNPGFFGRGLLSALSYLGETYQVISQAAVVHATA   KVEELLEQADYLYESGETEKLYQLLTQYKESGE   DAELLWRLARASRDVAQLSRTSEEKKLLVY   EALEYARRA/LÆKNESSFASHKWYAICLSD   GDYEGIGKAKIANAYIKEHIFEKARELANKDATS   HILMGIWCYTFAEMPWYQRRIA*NACLQLPG   FYPPYERALGYFFRARQVDPNFYSKNLLLIPG   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAAQLLTSFEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTSFEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTSFEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKKLAAPWLMKAKDYPAHTEED   KQIQTEAQALTAFSEEN   KTYLKLHNKAARAYIKATAFSEEN   KTYLKLHNKAARAYIKAARA	İ						
NITDTGFKLLUGUTINFUGGLHAGERARP   VPPRIPAPPAGGSKKRFQLERELAEQPW   VPRIPAPPAGGSKKRFQLERELAEQPW   VPRIPAPPAGGSKKRFQLERELAEQPW   VPRIPAGGIR   SSAGSARKLQVMALAARUWRLLPFRGAAP   SSRIPAGTSGSRGHGGPCRRGGFWMGNTGFFRGLLISALSYLGFETYQVISQAAVVHATA   KVEELEQADYLYESGETEKLYQULTQYKRSE   DAELLWRILARASRDVAQLSRTSEEEKKLUV   EALEVARAALEKNESSSAGHKWYAICLSDV   GDYEGIKAKIANAYIKEHFEKAELINFKDAT   HIMMIGWCYTFAEMPWQRRIA-NACLQLPP   *PPPYEKALGIYFHRABQVDPNFYSKNLLLLQ   GDYEGIKAKIANAYIKEHFEKAELINFKDAT   HIMMIGWCYTFAEMPWQRRIA-NACLQLPP   *PPPYEKALGIYFHRABQVDPNFYSKNLLLLQ   KYLKLHNKKLAAPWLMKAADVPAHTEED   KQIQTEAAQLLTSFSEN   KQIQTEAAQLTSFSEN		1					7
	Ì	1				į	
PPLPAGPLR		1				ł	
790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGRGHGCPRRGEPMGNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEELEQADVI-YESGETEKLYQLTGYKESE DAELLWBLARARSDVAQLSRTSEEEKKLLVY EALEYAKRA/LEKNESSFASHKWYALCLSDV GDYEGIKAKIANAYIKEHFEKAIELNRKDATS HILMGIWCYTFAEMPWYQRRIA-NACLQLP **PPPYEKALGIYPHRAEQUPPNFYSKNLLLIG KTYLKI-HNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN  791 2141 A 6434 3 1460 IALLIYDGLAWDDQGGLALLHISPSKLIL-QDS SGMSJYVMVRCTITRAFFSLLCHICQYSIGPQ **VNCPQQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAELARXYSSISINGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTIGIKQKQEPPIDLTSVPNLFTYNSS FNNAGLAPGITMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAQITHAIQYVVEFAKRITGFMELQNDQ ILLIKSGCLEVVLVRMCRAFNPLNNTVLFG KYGGMQMFKALGSDDLVNBAFDFAKNLCSL QLTEEEIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEEIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYFALQHYQKNHLDDEITAPAKNLCSL QLTEEGIALFSSAVLSPDRAWLEPRKVGLQ EKNYSHLDDEILAFTAPALG AGAGCEVTALMCDLRRAAGGMMDLAYV CEWEKWSKSTHCPSVFLACAWSCNNLLATM DLRSDQDLTLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDLTLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDDLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDDLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDQLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDQLTRUBHLDTEHPWDLHSIPSEHH DLRSDQDQLTRUBHLDTITHENDLRSCREWVSEKCRIDT DLRSLFMRCTIDLNRKDKPFAITHLKFLARD MSGQVLTANSESSYGSLVGGGPHLAWRLLAHP FQUSPVYIGKQCPTILKWBLISATNDLDRVSA VALPKLPISLTINTDLKVASDTQYFYGGLALA FFIGOSVHIVYSLSLQDPN	i	ı		l		i	
GSRLPAGTSĞSRĞHCGPCRRRGFEVMGNPCT FKRGLLISALSYLGFETYQVISQAAVVHATA KVEELIEQADVI.YESGETEKL.YQLLTQYKESE DAELLWRLARASRDVAQISRTSEEEKLLVY EALEYAKRA/LEKNESSFSAHKWYAICLSDV GDYEĞIKAKIANAYITKEHFEKAIELNFKDATS HILMGIWCYTTEAMPWYQRRIAYNACLQLPP *FPPYEKALĞIYFHRAEQVDPNFYSKNLLLLĞ KTYLKLHNIKLLAAFWIMKAKDYPAHTEED KQQTEAAQLLTSTSEKN  791 2141 A 6434 3 1460 IALLIYDĞLAWDDQĞĞLALLHISPSKLIL-QDS SĞMSYYVMYĞCTTRAFFKSLCHİCQYSİĞÇĞ *VYCCPĞQDACKE*KSTAN*GĞ*RE**PQVLFF AFLSYPAVKFRMSKKQRDLAVOKİĞQĞ RLĞÇRQQQŞĞBAEALARVYSSSİSNĞLSNIN NETSĞTYANĞSVDLPKSSĞYLAVQKİĞQĞ RLĞÇRQQQŞĞBAEALARVYSSSİSNĞLSNIN NETSĞTYANĞSVDLPKSSĞYLAVQKİĞQĞ RLĞÇRQQQŞĞBAEALARVYSSSİSNĞLSNIN NETSĞTYANĞSVDLPKSSĞYLAVÇKİĞQĞ RLĞÇRQQQŞĞBAEALARVYSSSİSNĞLSNIN NETSĞTYANĞYDLPKSSÇYLÖKLĞÇNİĞĞ LILLKSĞCLEVVLVIMCKARPININTVLFĞ KYĞĞMQMFKALĞSDDLVIRAFDFAKNLCSL QLTEETALFSSAVLISPPRAVINÇSİQPSI QQCAQİTHAİQVVVEŞAKRTİĞFMELÇÖNDĞ LILLKSĞCLEVVLVIMCKARPININTVLFĞ KYĞĞMQMFKALĞSDDLVIRAFDFAKNLCSL QLTEETALFSSAVLISPPRAVINÇSİQ LILLKSĞĞLEVVLVIMCKARPININTVLFFQ KYĞĞMQMFKALĞSDDLVIRAFDFAKNLCSL QLTEETALFSSAVLISPPRAVINĞLİY LÜĞLEBLALĞYĞMÇĞNİLİYENİNÜLLERNQRISH LÜĞLEKLĞYÖKĞĞĞYĞLİYENİNÜLLERNĞPRSH EŞÖÇĞMDLINDIVKLİSKDEVKAYLKÇKEPKH ĞAVVNISHİDƏL VEDIÇINKEĞRKİLİPPA LĞYĞKEĞKĞKİPPESTLİRNİDLLERNĞPRSH EŞÖÇĞMDLINDIVKLİSKÖEVKAYLKÇKEPKH ĞAVVNISHİDƏL LÜĞLEALĞĞÜĞĞĞĞAVÇÜNÜLKÜĞÇÜNÜL LYYTEĞĞĞAR ÇÜNÜLLİYÜL LÜĞLEALĞĞÜNĞĞA VÜNÜNÜLERPAAĞĞMMDLAYV CEWKEWĞKSTRİÇSVPLACAWƏĞRNILAFTM DERSÖNDÜNDÜNÜLKSKÖEVKAYLKÇEPKH ĞAVVNISHİDƏLEN ÇÜNÜNÜ ÜNÜNTÜLGÜNÜNÜNÜL LATLIYÜNĞĞĞYİLĞYÜNÜLĞĞĞİN VÜNÜNÜNÜLALHIVESĞĞASFĞÇEKFSRIVÇES PISLTLÇĞĞĞAMĞĞYLİLYÜNĞLÖYVISLÖQÜP SĞQVLUSTESILÇRERRAVALADİLFTĞĞĞNİ VATADĞSSASRYQPYKVÖNYEĞÇÜPÜL VATADĞSSASRYQPYKVÖNYEĞÇÜFÜL NÜĞLEL PÖĞĞÖYLÜNÜNÜLĞANDEN SAVALPKLIPSILĞINTÖLLEKARAPYSSAAPPPYÖ VALFAÇĞÖNLİLEN KILKISATNDLEN SAVALPKLIPSILTINTÜLKVASDTQFYYĞLĞLAL AFIÐĞOSYHİLYKİLSLÖTMAVYSSAAPPPYÖ	500			6000		1000	
FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEELEGADVLYSEGETEKULOTLOTYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/LEKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIKEHEKAIELINFKDATS HILMGIWCYTFAEMPWYQRRIA*NACLQLPP *PPPYPEKALGIYFHRAEQUPPNYSKNLLLIG KTYLKLHNKKLAAFWIMKAKDYPAHTEED KQIQTEAQLLSTSEKN  791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGLALLHISPSKLILI,*ODS SGMSYYMWRCTITRAFFKSLLCHICQYSIGPQ *VYLCPQQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQGNPQ RQCQRQQQSGEAELARXYSSSISNGLSNLN NETSGTYANGSVIDLFKSEGYYNVVSGOPSP DQSGLDMTIGHIQKQCEPIVDLTSVPNLFTYNSS PNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELIGILAWQHTTYSEIKAYSSSISNGLSNLN NETSGTYANGSVIDLFKSEGYYNVVSGOPSP DQSGLDMTIGHIQKQCEPIVDLTSVPNLFTYNSS PNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELIGILAWQHTTYSEIKAY QCAJOTHAIQVVVEFAKRITGFMELCQNDQ ILLKSGCLEVVLVRMCRAFPPLNULTPV TMEELIGILAWQHTTYSEIKALKAIFITTA VCNLHGEKLQVFKQSHEEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCSSMRLFLWNAVLLFV TSLIGALIPEPPVILEVLQKPFICHRKTKGGDL MLYHYBGYLEKDGSLFHSTIKHNNGQPIWFT LGILEALKGWGPG*KDMCVGEKKLIPPA LGYGKEGKGKPPSTLIFINLSKEDEVKAYLKKEPEKH GAVVNESHEIDALVEIDFINKEDEDKDGFISAR LGYGKEGKGKPPSTLIFINLSKEDEVKAYLKKEPEKH GAVVNESHEIDALVEIDFINKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRIKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKTHCPSVPLACAWSCRNLLAFITM DLRSDQQDLTRMHHLDTEHPVDLHSIPSEHH EAITCLEWDQSGFFGFLFSRWFTQGKCWS MGVSTLANISWESSYGSLVSGGPFLWALS MLYHNGVKLALHVERSGASSFGEKFSKIVFS PSLTLFGGRAMAGGWIALTVSGLVTVSLLQUP SQQULTSTESLCRLRRVALADIAFTGGGNI VVATADGSSAISPVGFYSKCKSTHCPSVSKCKREGIPVNNI FQQISPVVGBKQCFTLLKWRLSTDLDRXSA VALPKLIPSLTNTDLKVASDTQFYPGGLAL AFHDGSVHIVYRSLAQPHYVSSAAPRPVD	790	2140	A	6380	76	1059	,
KVEEILEQADYLYESGETEKLYQLITOYKESE DAELLWALARASRDVAGKSTSEEEKKLLVY EALEYAKRA/LEKNESSFASHK WYAICLSDV GDYSGIKAKIANAYIIKEHFEKAELINFOATIS HILMGIWCYTAEMPWYQRIA*NACLQLPP *FPPYEKALGYFHRABQVDPNTYSKILLLG KTYLKLINKKILAAPWLMKAKDVPAHTEED KQIQTEAQLLTISFEKN  791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGLALHISFSKILL;*QDS SGMS/YWNRCTITRAFFKSLLCHICQYSIGPQ *VTCPGQDACKE*KSTANGG*REP*PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEALARVYSSSISINGLSNIN NETSGTYANGSVIDLYSGGYSIVVSGQPSP DQSGLDMTNGIKQIKQPPIYDLTSVPNLFTYNSS PNMQCLAFGTIMTFEIDRAJNIKSSHLETCQY TMEELHQLAWQTHTYEEKAYQSKSREALW QQCAIQTHAIQYVVEFAKRITGFMCLQNDQ ILLIKSGCLEVVLVMCKAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEFALHSSAVLISPRAWLEPRKVQKLQ EKIYFALQHYQKNHLDDETLAKLJAKPTITA VCNLHGEKLQVFKQSHPEIVNTLFPLYKEF KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEFALRSSAVLISPRAWLEPRKVQKLQ EKIYFALQHYQKNHLDDETLAKLJAKPTITA VCNLHGEKLQVFKQSHPEIVNTLFPLYKEF KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEFALRSSAVLISPRAWLEPRKVQKLQ EKIYFALQHYQKSHPLAKNLTHAV VCNLHGEKLQVFKQSHPEIVNTLFPLYKEF KYGGMQMFKALGSDDLVNEAFPLYKEFK LGYGKEGKGKIPPESTLIFNIDLLERNGPSH LGYGKEGKGKIPPESTLIFNIDLLERNGPSH LGYGKEGKGKIPPESTLIFNIDLLERNGPSH LGYGKEGKGKIPPESTLIFNIDLLERNGPSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKGGFISAR EFTYKHDBL  793 2143 A 6446 3201 152 PRIKRILVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRPPAGGGMMLAYV CEWBEWSKSTICPSVPLACAWSCRNLJAFTM DLRSDDQDLTRMIHLIDTEHPWDLHSIPSEHH EAITCLLEWDQSGPFGFLSRWPTOQIKICWS MGVSTLANSWESSVOSLVEGGPHL WALS WLHNGVKLALHVEKSGASSFGERYSIKVFS PSITLF1GGNAMEGWIAVTVSGLVTVSLLQVP SGQVLTSTESLCRLRARVALLADLAFTGGGN VVATAGGSSASPVQFYKCVSVYSEKCRDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRECHDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRECHDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRECHDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSQVULCASSQTSSIVECWSVSEKCRECHDT DDLPSLFMRCTTDLNRKDKFPAITHLKFLARD MS					ļ		
DAELLWRI.ARASRDVAQUISITSEERKILLVY EALEYARA/LEKNESSASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS IHLMGIWCYTFAEMPWYQRILA'NACLQLPP *PFPYEKAIGYFHRAEQVDPNFYSKOLLLLG KTYLKLENKKI.AAFWLMKAKDVPAHTEED KQIQTEAQAQLTISTSEKU KQIQTEAQAQLTISTSEKU (NICTUAL CONTROL TO SGMSYVMYRCTITRAFFKSLLCHICQYSIGPQ *VYNCPGQDACKE*KSTANGG*RE**PQVLFF AFLSDRAVKFGRMSKKQRDSLYAEQVKHQQ RLQEQRQQOSGAEALARYSSSISNIGSNIN NETSGTYANGSVDLPKSEGYYNVVSGOPSP DQSGLDMTIGIKQIKQEPYDLTSVPNLFTYNSS FNNGQLAPGITIMTEDIRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSEALW QQCAIQTHAIQYVVETRAKRITGFMELCQNDQ ILLILKSGCLEVVLVBMCRAFNPLNNTVLFEG KYGMQMGKALGSDDLVBAEPDFAKNLCSL QLTEEFIALPSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDAWLIEPRKVQKLQ EKIYFALQRIVQKNHILDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHFDESTDETNGPPSH ESSGEMDLNDDWKLISKDEVAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR ESFQEMDLNDDWKLISKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRIKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRPPAAGGMMLAYV CEWBKWSKSTTCPSVLACAWSCRNLLAFITM DLRSDDQDLTRMHILDTEHFWDLISIPSEHH EAITCLEWDQOSGPFGELFSRWPTOGIKCWS MGVSTLAINSWESSVGSLVEGGPHLWALSI WLHNGVKLALHVEKSGASSFGELFSRNVFTS PSS.TLPGGGNAMEGWUAVVSGLVTVSLLQVP SGQVLUTSTESLCRLRARVALLADLAFTGGGNI VVATADGSSASPVQFYKVSVSEKCRDDT DILPSLAMRCTTDLINRIDKEPRATTGGNALGVIIAVFYSGLAPTTGGGNI VVATADGSSASPVQFYKVSVSEKCREDT DILPSLAMRCTTDLINRIDKEPRATTGGNALGVIIAVFYSGLAPTTGGGNI VVATADGSSASPVQFYKVSVSEKCREDT DILPSLAMRCTTDLINRIDKERSATTCDLORVSA VALPKDSSVIVHKHLSLSTNDLDENSA VALPKDSSVIVHKHLSLSTNDLDENSA VALPKDSSVIVHKHLSLSTNDLDENSA VALPKDSSVIVHKHLSLSTNDLDENSA VALPKDSSVIVHKHLSLSTNDLDENSA							
### PART   PART	1		1	1			
GOYEGIKAKIANATIKEHFIFKAJELNIPKDATS   HHMGIWCYTFAEMPWYQRIA*NACLQLPP	1		l				
### ### ##############################	1		!	1			
#FPPYEKALGYPHRAEQUDPNFYSKNLLLLG KTYLKLHNKKLARPHAEODDRAKKARDYPAHTEED KQIQTEAAQLLTSFSEKN  791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGLALLHSPSKLLCHICQYSIGPQ VTVCPGQDACKE*KSTAN*GG*RE**PPQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEALLARVYSSSISNGLSNLN NETSGTYANGSVIDLFKSEGYYNVVSGQPSP DQSGLDMTGKQKQKQEPTYDLTSVPNLFTYNSS FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMELHQLAWQTHTYLFTYNSS FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMELHQLAWQTHTYUNFAPTPAKNLCSL QLTEEEIALPSSAVLISPDRAWLIEPRKVQKLQ EKTYFALQHVIQKNHUNFAPTPAKNLCSL QLTEEEIALPSSAVLISPDRAWLIEPRKVQKLQ EKTYFALQHVIQKNHUNFAPTPAKNLCSL QLTEEEIALPSSAVLISPDRAWLIEPRKVQKLQ EKTYFALQHVIQKNHUNFAPTPAKNLCSL QLTEEEIALPSSAVLISPDRAWLIEPRKVQKLQ EKTYFALQHVIQKNHUNFAPTPAKNLCSL NUMLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGITRCFCRDFTPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYSTYEJLEKDGSLTHSTHKHINNGQPWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKPPESTLIFNIDLLEIRNGPRSH ESTQEMDINDDWKLSKDEVKAYLKKEFEKH GAVVNESHEIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFIM DLRSDDQDLTRMIHLDTEIPPWDLHSIPSEHH EATTCLLEVDQSGFFGFIFSRWPTGQIKCWS MGVSTLANISWESSVGSLIVEGGPH WALS. WLHNGVKLALHVEKSGASSFGEKFSR/VKFS PSSTLTEGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLICTSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPFVGKYSSVVSSKECRIDT DILPSLFMRCTTDLNRKDKFPATHLKFLARD MSEQVILCASSQTSSIVECWSLKKEGLPVNNI FQQISSVVGDKQFTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTOTYPYGLGLAL AFHDGSVMIVHRLSLQTMAVTYSSLATRPYD	1		1	1		Ī	
KTYLKLIHNKKLAAFWLMKAKDYPAHTEED							1
RQIQTEAAQLLTSFSEKN	1						
791 2141 A 6434 3 1460 IALLTVDGLAWDDQGGLALLHISPSKILI-QDS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ  *VTICPQQDACK#-VENTAN*GG*RE**PQVLFF ARLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNIN. NETSGTYANGSVIDLYKSEGYYNVVSGQPSP DQSGLDMTIGIKQIKQEPIYDLTSVPNLFTYSS FNNGQLAPGITMTEIDRIAQNIKSH.ETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLKSGGLEVVLVMCRAFNPLINTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEFIALPSSAVLISPDAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTTRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALPEPEVKIEVLQKPFICHRKTKGGDL MLVYEGYLEKGGSLFISTHKHNNGQPIWFT LGILEALKGWGPGA*KDMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVTFEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRAAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGLFSRWPTGQIKCWS MGVSTLANSWESSVGSLVEGGPHLWALS WHINGVKLALHVEKSGASSFGEKFSRIVKFS PSLTLFGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLJTSTESLCRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFATHLKFLARD MSGQVLJCASSGTSSVECWSLKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTOFFYCIGLAL AFHDOSVHVHRILSLQTMAVFYSSAAAPRPVD	1				ł		
SGMS/YVMYCTITRAFKSLLCHICQYSIGPQ *VT\CPQQDACKE*KSTAN'GG*FE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARYYSSISIGLSINLN NETSGTYANGSVIDLPKSEGYYNVVSQPSP DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTY\SS FN\\GQLAPGIT\MTEIDIAQ\\RikSHLETCQY T\\MEELHQLAWQTHITYEIKAYQSKS\REAL\\QQCAIQ\THAIQ\\YVVEFAK\RITGFMELCQ\\RDQ\\LILLK\SGCLEVYLV\RMC\RAF\\PPL\\RNT\VLFG\\K\G\\K\G\\RMC\RAF\\RPPL\\RNT\VLFG\\K\G\\K\G\\RMC\RAF\\RPPL\\RNT\VLFG\\K\G\\K\G\\RMC\RAF\\RPPL\\RNT\VLF\\RPPL\\CA\\K\G\\RMC\RAF\\RPPL\\RM\\LIP\RY\\\G\\K\G\\RMC\RAF\\RPPL\\RM\\LIP\RY\\\G\\K\G\\RMC\\RPPL\\RM\\LIP\RY\\G\\K\G\\RM\\RPP\\RM\\RM\\RPP\\RM\\RM\\RPP\\RM\\RM		ļ	ļ.,				
**VTNCPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQEPIYDLTSVPNLFTYNSS FNNGQLAPGITIMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQTHAIQVVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALPSSAVLISPDRAWLEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 192 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALPEFEVKIEVLQKFFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGI EALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEKKGKIPPESTLIFNIDLLEIRNGFRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRPPAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNILAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLLEWDQSGFPGHLSRWFTGQKKCWS MGVSTLANSWESSVGSLVEGGPHLWALSI WLHINGVKLALHVEKSGASSFGEKFSRVKFS PISLTLFIGGNAMEGWHAVTVSGLVTVSLLQIP SGQVLTSTNESLCRLRAVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRREGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VVALPKLPISLINTIDLKVASDTQFYFGIGLAL AFHDGSVHIVHRSLQTMAVYSSAAPRPVD	791	2141	A	6434	1 3	1460	
AFLSNPAVKFGRMSKKQRDSLYABVQČHQQ RLQEQRQQQSGAEALARVYSSSISNGLSNIN NETSGTYANGSVIDLPKSEGYYNVVSQQPSP DQSGLDMT\GIKQIK\QEPYPDLTSVPNLFTY\SS FNN\QLAPGIT\mTeIDRIAQ\NIKSHLETCQY T\MEELHQLAW\THTYEIKA\QS\SSRFAL\W\QQCAIQ\THAIQ\VV\EFAKRITGFMEL.Q\NDQ\ ILLILKSGCLEVYLV\MCRAF\PPL\N\TVLFIG\ K\YGGM\QMF\KALGS\DDL\W\EAF\PFAK\\LCS\L\Q\EK\YFAL\Q\VV\EAF\PPL\N\TVLFIG\ K\YGGM\QMF\KALGS\DDL\W\EAF\PFAK\\LCS\L\Q\EK\YFAL\Q\V\EX\H\LD\DT\L\A\LI\A\LI\A\LI\A\LI\A\LI\A\LI\A\LI\A							
RLQEORQQOSGEAEALARVYSSSISNÖLSMÄN NETSGTYANGSVIDLPKSEGYYNVVSGOPSP DQSGLDMTIGIKQIKQEPIYDLTSVPNLFTTYSS FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDEILAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILEALKGWGFGA*KDMCVGEKRKLIIPPA LGYGKEGKGKPPESTLIFNDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRFRAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EATTCLEWDQSGFPGFLFSRWPTGQIKCWS MGVSTLANISWESSVGSLIVGGGPHLWALS WLHWIGVKLALHVEKSGASSFGEKFSRIVKFS PSLTLFGGNAMEGWIAVTVSGLVTVSLLQP SGQVLITSTEGLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVILCASSQTSIVECWSURKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTINTDLKVASDTQFYPGLGLAL AFHDGSVHIVHILSLQTIMAVFYSSAAPRPVD		İ				}	
NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQEPIYDLTSVPNLFTYYSS FNNRQLAPGITMTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEILALFSSAVLISPDRAWLEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHHNNGQPIWFT LGILEALKGWGPGA*KJMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESSQEMDLNDDWKLSKDEVKAYLKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRIKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKICWS MGVSTLAINSWESSVGSLVEGGFHLWALS; WLHNGVKLALHVEKSGASSFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVLITSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVOFYKVCVSVVSEKCRIDT DILPSLFMCTTDLINKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLREGGLPVNNI FQQISPVVGDKQPTILK WRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTOFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1	i				
DOSGLDMTIGIKQIKQEPIYDLTSVPNLFTYISS FNNGQLAPGITIMTEIDRIAQNIIKSHLETCQY TIMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLILKSGCLEVVLVMRCRAFINPLINITVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKYFALQHYIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKEVLQKPFICHRKTKGGDL MLVHYEGYTLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKPPFESTLIFNIDLLEIRNGFRSH ESSPQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHLDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKICWS MGVSTLAINSWEISSVGSLVEGGPHLWALSI WLHNIGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVYTVSLLQVP SGQVL\TSTVESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSSGQVLLCASSQTSSIVECWSLRKEGLPVINNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFFIDGSVHIVHRLSLQTIMAVFYSSAAPRPVD			١.				
FNNGQLAPGITMTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHITYEIIKAYQSKSRAALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMMFKALGSDDLVNEAFDFAKNICSL QLTEEBIAPSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGHEALKGWGPGA*KJMCVGEKRKLIIPPA LGYGKEGKGKUPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRIKRLVVTEEDGGARPEALGKIAPRTPAELG ARADGELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHLDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGHSRWPTGQIK/CWS MGVSTLANSWE\SSVGSL\VEGGPHLWALS\ WLHNIGVKLALHVEKSGASSFGEKFSRVYKFS P\SITLF\GGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSGQVLLTSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSGQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1		ļ				
TMEELHQLAWOTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELQNDQ ILLIKSGCLEVVLVRMCRAFRPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKLEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEIPWDLHSIPSEHH EAITCLLEWDQSGPPGFLFSRWPTGQIKICWS MGVSTLAINSWESSVGSLVEGGPHLWALS, WLHNGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSANSPVQFYKVCVSVVSEKCRIDT DILPSLEMRCTTDLINRKDKPPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPYNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD			i	:			
QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ IILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTK.GDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGGLEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLLAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLAINSWESSVGSLVEGGPHLWALS\ WLHNNGYKLALHVEKSGASSFGEKFSRVVKFS PSLTLFIGGNAMEGWIAVTVSGLVTVSLLQIP SGQVL/TSTTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DLPSLEPMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	Į.	ĺ	1				
ILLKSĞCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFRALGSDDU VNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRA WLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK NPDCATACK NPDCATACK SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYVEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKEPPESTLIFINDLLEIRNGFRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHLLDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLANISWESSVGSLVEGGPHL WALSI WLHNGVKLALHVEKSGASSFGEKFSR\VKFS PSLTLF\GGNAME\GWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDIT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD			1		ĺ		
KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKUPFESTLIFNIDLLEIRNGFRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLAINSWEISSVGSLVEGGPHL WALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	ļ	Ī	l				,
QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKTYFALQHVIQKNHLIDDETLAKLIAKIPTITIA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGPFGFLFSRWPTGQIK/CWS MGVSTLA/NSWE/SSVGSL/VEGGPHLWALS/ WLH/NGVKLALHVEKSGASSFGEKFSR/VKFS P/SSLTLF/GGNAMEGWIAVTVSGLVTVSLLQ/P SGQVL/TSTLESLCRLRARVALADIAFTGGONI VVATADGSSA/SPVQ/FYK/CVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLP/NNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	Ι΄.						
PST 2142 A 6440 92 781 SRGTPRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGK/PPESTLIFNDLLEIRNAFSHH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK/CWS MGVSTLA/NSWESSYGSL/VEGGPHL WALSI WLHWGVKLALHVEKSGASSFGEKFSR/VKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVL/TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA/SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFFYPGLGLAL AFHDGSVHTVHRLSLQTMAVFYSSAAPRPVD		ļ.					
VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK  792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKVCWS MGVSTLAINSWESSVGSLIVEGGPHLWALSI WLHINGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVTVSLLQVP SGQVLITSTVESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNITDLKVASDTOFFYPGLGLAL AFFIDGSVHIVHRLSJQTMAVFYSSAAPRPVD			İ	ļ			1 -
792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGFGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKVCWS MGVSTLAINSWESSVGSLVEGGPHLWALSI WLHINGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLVTSTVESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTIMAVFYSSAAPRPVD	İ		ļ				
792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWESSVGSLVEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSRVKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVILCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK/CWS MGVSTLA'NSWE'SSVGSLVEGGPHLWALSI WLHNIGVKLALHVEKSGASSFGEKFSRIVKFS PSLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	702	2142				701	
MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLANSWE\SSVGSL\VEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGL\VTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLIKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	192	2142	A	6440	92	/81	l _ • - · · • · · · · · · · · · · · · · · · · ·
LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1	ĺ				
LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACA WSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSR WPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHL WALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1					
ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACA WSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWESSVGSLVEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		İ	l				
GAVVNESHHDAL VEDIFDKEDEDKDGFISAR EFTYKHDEL  793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWESSVGSLVEGGPHL WALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\PSGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1		i				
793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACA WSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSR WPTGQIK\CWS MGVSTLA\NS\WE\SSVGSL\VEGGPHL WALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDL\NRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSR WPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	702	101/5	<u> </u>	-	2201	1.50	
CEWEK WSKSTHCPSVPLACA WSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSR WPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHL WALS\ WLHINGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	793	2143	Α .	6446	3201	152	
DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD					1		-
EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD					İ		
MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1	1					
WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						EAITC/LEWDQSGFPGFLFSRWPTGQIK/CWS
P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						
SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	İ		i				•
AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
				'	•		
EPAMKRPRTAGPAVHLKAMOLSWTSLALVG							
		L					<u>EPAMKRPRTAGPAVHLKAMQLSWTSLAL</u> VG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	согтеspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	1	peptide		/=possible nucleotide deletion, \-possible
<u> </u>			ļ	sequence		nucleotide insertion
	•			1		IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
1		ł	1	i		LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
1						HEEYTRQTAALQQVLSTRILAMKASLCKLSP
1						CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
		l				PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF
1		l			ĺ	VLDMNTLQALQQLLQWVGDFVLYLLASLPN
		•	1		<b>,</b>	QPCPTSEPCPTSEPSPTSEPSSP*SLC\G
1		ì	ļ		İ	SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
			1		i	KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
						RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
						PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
			]			LQLDGLARAPGQPKIDHLRRLHLGACPTEEC KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
		ļ				NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA
1		1				APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
		ļ	) .	l	ļ	RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT
						PRSLDHLHPEDRP
794	2144	A	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
1	~	1	01,70	410	303	GSIEPRDLRLQ*AVITPL\TPAWVTQ
795	2145	A	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
		١.,	0.55	333	1027	SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP
		1				FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
1		i	1	<b>'</b>		DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN
		l				STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
1		l		,		HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
						DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
						DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
			1.			PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM
i i						KKMNKLQNISEELQRNISLQLMSNMNISNKIR
				1		NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR
			.			WIWHKDSCYFLSDDVQTWQESKMACAAQN
					l	ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
1						DS/YSWYESG*YNQ\PSAWVIRNAPDLNNMY
1 1			i i		ŀ	CGYINRLYVQYYHCTYKQRMICEKMANPVQ
100						LGSTYFREA
797	2147	A	6507	1	881	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
[						APRAASAQLEERMRDPHPGMTLQEGDCRGS
						QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
						ESALAKLLLTCCSALRPRATQARGSSRLLVAS
]					·	WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
		•	[	ĺ		AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
.				l		TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
				}		ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM
	l		_	ļ		DMLKALFRTLQAMLLGVWILLLASLTPLWL
				ļ		/SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
	ľ		!	ļ		STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
					ĺ	RAASAQLEERMRDPHPGMTLQEGDCRGSQT
					İ	VSLTMGTADSDEMAPEAPQHTHIDVHIHQES
						ALAKLLLTCCSALRPRATQARGSSRLLVASW
	1	ľ		l	l	VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
	1				1	WTGAVAVLAGAAAFIYEKRGGTYWALLRTL
j			l		ł	LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
	ĺ	- (	ĺ	ſ	. 1	RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM
1		•		[	. }	LKALFRTLQAMLLGVWILLLASLTPLWLYC
798	2148	A	6528	912	2287	WRMFPTKGVSP VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
'~	2170	^	0320	712	2201	RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
	l		ļ	l		LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG
<b></b>			1			DIATA ATMENDICATE I ASSEDILI I MALOREO

NO. of   No. of   N														
mucleotide sequence   unace	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,							
Sociation   Soci			hod											
unice   unic														
union and residue of peptide peptide sequence   p	1													
### mino acid residue of peptide requence peptide sequence peptide sequence peptide sequence peptide sequence per sequence peptide sequence per sequ	1 .	uence		1										
residue of peptide sequence	uence	t		914										
### Popsible nucleotide deteion, impossible nucleotide insertion uncleotide insertion uncleotide insertion in the property of	i	ļ	ł	l										
	1	Į.				sequence								
### STATES OF PRAISE OF PR	1	ĺ	ļ											
FPPAHENAFIVIASSLOHMLITCILWILITKK HTVSQEDOLSLAGAPQPRIKSRTSVLRIV MYRWELSSNGPGRGVLGIGIGGNSLRIV GONLGI-HCVWVVWETGF*KRWILOMGEF GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF STIVTGGGVG-IHVTEILPFEVKAAOPILL GPSLPQRQGREHIVVILAAPACAPFIDR*WEP RERPSPFELGRGSPT-ISYPASCRVIRQPFD RKSYSWKORLFINFISFTSALAVYRHDNYC EGAVYTIFALETYVVITMAFHMYAWDF GNKELLITSQPEKR?  799 2149 A 6529 1 874 FPFPGRNFTEINGSVSLLALACDLGWCEDWS CLVQGGDLVDVVQTNHGEDEAGDTDEV DEARCKESQQEAQENLREDLCLESFAKDKIL QHEGSRENHEETRIKQAALDGFHZGGQLTA VILHPSKQCQGEGGEQRGARTHHWRGW EGRRVRLRPSGKLAADQPVRKLGGTFST ELPGLQPHAPTHTAAATPTYSPADTINPPY RWKCFLVFERRTGLCRETKKACPKFPFPL GLPDDFTGFVTHHAPVSSTGASGCRRAEP FPFRENCHENTEN EGRRVRLRPSGKLAADQPVRKLGGTFST ELPGLQPHAPTHTAAATPTYSPADTINPPY RWKCFLVFERRTGLCRETKKACPKFPFPL GLPDDFTGFVTHHAPVSSTGASGCRRAEP FTASKNGTVKFCTSNESTFHKTYYTPPOVG ETHLCFLVRRAFSALTAVERSKCVSHEALKS VDVQTHFRLEAQGRSGRAEDLNTVAYWSV GEALLLLVVSIGQVFLLKSFFSDKRTTTTRVGS STAGNFFRENCHSCHSTTHKTYYTPPOVG ETHLCFLVRRAFVSALTQMESACVSHEALKS VDVQTHFRLEAQGRSRAEDLNTVAYWSV GEALLLLVVSIGQVFLLKSFFSDKRTTTTTRVGS STAGNFFRENCHSCHSCHSCHSTFTHKTYYTPPOVG ETHLCFLVRRAFVSALTQMESSCSTLCE GASTAMAVAVGVFRUSTGSCHSTTHKTYYTPPOVG ETHLCFLVRRAFVSALTQMESSCSTLCT RKIFLKRMSIRESLKERGVDMAALDGFBWSQP MKRLTLGNTTSSVILTNYMDTDYYGEIGIGTP POTFKVVFOTGSSTVWYPSSKCSSLTAGAVY HKLFDASDSSSYKINGTELTLRYSTGTVSGFL SADYLPGSVSSKLSTLAHAMYPPPTOFTL VALGATPIRKYTTEDRORMPHGFALAR COCKARAPTER PROTECTISCL SADYLPGSSSSKLSTLAHAMYPPTOTTL VALGATPIRKYTTEDRORMPHGFALAR VOTGAKDGRAEGURGVFPVFNOMWIVKG UPGSQCKYSTGSTSIEKLMERMPHY GSVYTKGNPGSGGRKVELVGESTYCTSND DCYGTWGRAPQCUIPNTCTPTNVENGHLVST VOTGAKDRAGASSRREPEPGGPARATLPTCCGGL LAVVYLALPVAWQCNAFEWLFFARPTNL TIDEFEPFIGTTNYCCRGVGPPOVLAGRENFHY GSVYTTKGNPGSGGRKVELVGESTYCTSND DCYGTWGRAPQCUIPNTCTPTNVENGHLVST UPGAKDPTVCDEGGGPPOVLGRASSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOVLAARSMCTPQG DWSPAATCTCVSCODPPOV	<u> </u>				sequence	ļ								
HITVSGENGLSLAGAPROPRIKSRITSVLRIRV MYRWELSSINGNPGROVLGLIGIGGINSLRV GQNLGL-HCVWVVWETGEYKRWRLQMGE; GVASRQ-VRINSVRGVLGHISSAPPMYMGFF SPTVFGGGVGG-LHVTHTLHPPEVTAAGPILL GPSLPGNGGREHIVVLLAAPACAPHDEV-WEP REIRPSP-ELGLAGEFTLSYPASCRVIRQPIP-D RKSYSWKQRLFINFISFTSASCRVIRQPIP-D RKSYSWKYCQRCFSSPYVIRGRRTICLUSS RKSIPLSSCRVIRQPIP-D RKSYSWKCQRCFSSPYVIRGRRTICLUSS RKSIPLSSCRVIRQPIP-D RKSYSVCRCGRCFSSPYVIRGRRTICLUSS RKSIPLSSCRVIRQPIP-D RKSYSWKCQRCFTSS														
MVRWELSNGNPGROVIGLICLGLGHCHNLRVV   GONLGL-HCWVWETGe*RKUJMGHE*  GVASRRQ*VRNSVRGLVCHNSSAPPMYMGL    GVASRRQ*VRNSVRGLUCHNSSAPPMYMGL    GVASRRQ*VRNSVRGLUCHNSSAPPMYMGL    GRSLPQRQGREHIVVILAAPACAPPHR*WPPP    RERPSP*PLGLRGEPTLSYPASCRVRQPPPP    RKSYSWKQRLFINNFISFFSALAVYRRHMYC    EACVYTIFALLEYTVU.TNMAFHMTAWMYC    EACVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVYTIFALLEYTVU.TNMAFHMTAWMYC    GAVRULAYPSSULLALACDLGWEEDWS    CLLVQGGGLVDVVQTNHGEDEAGGDTDSV    DEARCKESQCAQENIREDLCLESPANKHI    GURGERRHEETTRIKQAALDGEPLGGQLTA    VILHFSKQQQGEGGRORGARTHHWRGW    EKORRYRLPPPSKGLRADQPVRLGOFTEST    ELFGLQPHAFTHTAPATPTYSFADTIFNYC    GRYCHLPPERTRIQCERSTRKACPYKFRPPL    GLPGDFTOPVTHHAPPVSFTGASGGERRAEP    GAVSYAHASATK    GRYCHLPPERTRIQCERSTRKACPYKFRPPL    GLPGDFTOPVTHHAPPVSFTGASGGERRAEP    TASKNGTYKFCFSNETSTTHRKYTYDFDQVG    ETHLCFLVRDRYSALTQMESACVSHELBLX    VIDNYGTHFILREAGGRSRAEDLINTWAYWSV    GEALILLVYSIQQVFLLKSFTSDKRTTTTXVS    GHYDVDCRLEDPDGKVLYKEMKKQYDSTT    TASKNGTYKFCFSNETSTTHRKYTYPDPQVG    ETHLCFLVRDRYSALTQMESACVSHELBLX    VIDNYGTHFILREAGGRSRAEDLINTWAYWSV    GEALILLVSIGQVFLLKSFTSDKRTTTTXVS    GEALILLVSIGQVFLLKSFTSDKRTTTTXVS    STORE     GHYDVDCRLEDPDGKVLYKEMKKQYDSTT    KRIFTKAMPSIRESLEKRGVDMARLGPEWSQP    MKRLTLGNTTSVLTNYMDTQYYGEIGIFP    PQTFEKVVPDGSSNWVPSSKCSRLYTACVY    HKLFDASDSSSYKHNGTELTLRVSTGTVSGGIF    PQTFEKVVPDGSSNWVPSSKCSRLYTACVY    HKLFDASDSSSYKHNGTELTLRVSTGTVSGGIF    PQTFEKVVPDGSSNSWVPSSKCSRLYTACVY    HKLFDASDSSSYKHNGTELTLRVSTGTVSGGIF    PQTFEKVVPDGSSNSWVPSSKCSRLYTACVY    HKLFDASDSSSYKHNGTELTLRVSTGTVSGGIF    SQDITTYGGTVTVDFGVFTEMPHALEP   DGVVGMGFEQAIGRYTPIDNISQOVLKE    SQDITTYGGTVTVDFGVFTEMPHALEP   DGVVGMGFEQAIGRYTPIDNISQOVLKE    VSFYYNRDSSNSSVCSSSTSLTACAGGR   SQDITTYGGTVTVDFGVFTEMPHALEP   DGVVGMGFEQAIGRYTPIDNISQOVLKE    GSVYTYRCNGSGGRKVELVGERSYCCSDS    VWTGARDRCRRSKKSITALHAMPOKANKTPIV    DRETTIFOTYLNTVECPGOYSGRPSILCLNS    VWTGARDRCRRSKKSNPPDPVGMAFMICHINGK    UKANTYVCCPGFTSVT		l												
GQN.G.: HCVWVVWETGE*RRWRLQMGEF GVASRQ@VGN:VRGLVCHNSSAPPMYMGEF SPTVFGGGVGG*LHVTELLPFEVEAAGPHOR*WED RETRRSP*BLGL.RGEPTLSYPASCRVTRQPTP*D RESPONDED TO THE STATE OF T		l	1				1							
GVASRRQ-VRNSVRGLUCHNSSAPPMTMEP   SPIVEGGVGG-LHVTILLIPPEVBAG0PILL   GPSLPQRQGREHIVVILAAPACAPPHR* WEP   RERPSP**ELG RGEPTLSYPAS-CRVRQPPPD    RKSYSWKQRLFINPISFFSALAVYFRHNMYC    EAGVYTTFALLEYTVVLTNMAFHMTAWDYC    GNELLITSQPEEKF    GNELLITSQPEEKF    GNELLITSQPEEKF    FPFPQRINFIERSGSVSLLACCDLGWCEDWS    CLVQGGGDLVDVVQTNHGEDEAGGDTDSV    DEARCKESQQEAGENLREDLCLESPAKDKIL,   QIEGSERHEFITKQAALDGEPGGGLTA    VHLHPSKEQQQGGGERQRGARTHHWROW    EKGRYRLRPPSKLRADQPVRKLGGFTBYT    ELPGLQPHAFTPHTAPATPTYSPADTTPMY    RWKCPLPVEPRTRQLCRETRKACHGGQLTA    VHLHPSKEQQQGGGERQRGARTHHWROW    EKGRYRLRPPSKLRADQPVRKLGGFTBYT    GLPGDPTOPVTHHAPPVSPTGASGQERRAEP    GAVSYAHASATK    GAVSYAHASATK    BOD 2150 A 6544 2 662 SAQRWAAVGGWGCRLIALLLLYPGPGGS    ETTELPDNAKQCFPEDLQGTKCTLEFQVITG    GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF    TASKNGTYKPCFSNEVSTIFTHKTVYFDROV    ETHLCHLWDROVSALTQMESSCSCREPQTESSU    GHALLVVSIGQVFLLSSFPSKACTSHEALKS    VIDYQTHFRLREAQGRSRAEDLNTRVAYWS    GEALLLVVSIGQVFLLSSFPSKACTSHEALKS    VIDYQTHFRLREAQGRSRAEDLNTRVAYWS    GRALLLVVSIGQVFLLSSFPSKACTSHEALKS    VIDYQTHFRLREAQGRSRAEDLNTRVAYWS    GRALLLVVSIGQVFLLSSFPSKRTTITKVS    GRALLLVVSIGQVFLLSSFPSKRTTITKVS    GRALLLVVSIGQVFLLSSFPSKRTTITKVS    GRALLLVVSIGQVFLLSSFPSKRTTITKVS    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDFMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDFMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDFMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDFMARLGFBWFLC    KRIFLKRMFSIRESLKERGVDFMARLGFBWFLC    KRIFLKRMFT    KRIFLKRMFT    KRIFLKRMFT    KRIF			1			ŀ								
SPTVFGGGVGG*LHVTFLLIPEVENAGPIPT* WED REIRRSP*ELGL.RGEPTLSYPASCR.VIRQPIP**	1	l	l	Į.			1 a. 7 a. GGDTDSV DEARCKESQQEAGENLREDLCLESFAKDKL QIEGGSREHEEITKQAALDGEPLGGGQLTA VHLHFISKQQQQEGGRQGAATHHWRCH EKGRYRLRPPSKLRADQPVRKLGGFTEST ELPGLQPHAFTPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHAFTPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHATPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHATPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHATPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHATPHTAPATPTYSPAPDTTPPY RWKCPLPVEPTRIKQLCRETRKAVLGGFTEST ELPGLQPHATTHAPPYSPTGASQCRRAEE GAVSYAHASATK   GLAVATAGRWCCRLLALLLLVPGPGGAG GAVSYAHASATK   GHANDAVLGARGA GAVENDAVLGARGA GAVAGRWCCRLLALLLLVPGPGGAG GAVSYAHASATK   GHANDAVLGARGA GAVAGRWCCRLLALLLLVPGPGAG GAVSHEALKS VIDYQTHERLREAQGRSRAEDLATRVAYWY GEITHLCHVRDRVSALTQMERKSCNSTFTTTRVGS GALLLLVVSIGQVFLLKSFPSOKATSTTTRVGS GALLLLVVSIGQVFLLKSFPSOKATSTTTRVGS GALLLLVVSIGQVFLLKSFPSOKATSTTTRVGS MKRILLTGATTSTTTVGGF KRILLVVSIGQVFLLKSFPSOKATSTTTRVGS MKRILLTGATTSTTTVGGF KRILLVVSIGQVFLKSFPSOKATSTTTRVGS MKRILLTGATTSTTVGGF KRILLVVSIGQVFLKED QVGGAGGAGATVFTEPDNISQQVLKED QVGGAGGAGATVFTEPDNISQQVLKED QVGGAGGAGATVFTEPDNISQQVLKED QVGGAGGAGAGATVFTEPDNISQQVLKED QVGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	ł						I
RERRSP'ELGLRGEPTLSYFASCRUNGPIPD RKSYSWKQRLININISISTSALAVYFRINMYC EAGVYTTAILEYTVVLTMMAFHMTAWWDF GNKELLITSOPEKER  799 2149 A 6529 1 874 PFFFQRINFIEHSGSVSLLALACDLGWCEDWS CCLVQGGGDLVDVVQTNHGEDEAGGDLTA VHLHPSKEQQGEQGERGRAGALTHWRGW EGRRVRLRPPSGKLRADQPVRKLGGPTEST ELGLQPHAPTPHTAPATTYSFAPDTPNPYV RWKCPLPVERTRQLCRERTRKACPPKPRPPL GLROPTIGPVTHTAPATTYSFAPDTPNPYV RWKCPLPVERTRQLCRERTRKACPPKPRPPL GLROPTIGPVTHAPPVSFTGASGGERRAP GAVSYAHASATK  800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLVPGPGGAS EITFELDDNAKQCFYEDIAQGTKCTLEFQVTIG GHYDVDCRLEDPDGKVLYKEMKKQYDSTF TASKNGTYKFCSNEFSTIFTNYVYDFDVQ EITHLCFLVRDRVSALTQMESACVSHEALKS VIDYQHTRLREAQGSRRADITNRVAYWSV GEALLLLVVSIGQVFLLKSFFSDKRTTTTYRGS TOMBORITARYAYWSV GEALLLLVVSIGQVFLLKSFFSDKRTTTTYRGS POWRRMPRWGLLLLLWGSCTFGLPTDTTF KRIFICKNPSTRESLKERGVDMARLGPEWSOP MKRLTLGNTTSSVILNYMDTQYYGEIGIGTP PQTFKVVPDTGSSNVVPSSKCSRLYTACVY HKLFDASDSSSYKINGTEILTRYSTGTVSGFL SQDITYGGTTYQMFGEVTEMPALPPMLAEP DGVVGMGFEQAGIGRVTPPINSIGQVLKED VFSFYYNRDSENSOSLOGQVLGGSDPQHYE GNPHYNLIKTGWQQMGKQSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLPDVVKCNGEPTLPPTIFLLGGKDTPLT SADYLPGESYSSKKLSTLAHHAMYPPPTPTL LAGATBURKYTEPDGRANDFFALAR  802 2152 A 6567 13 6147 MCLGRMGASSFRSPEPVGFPAGLFFCCGGSL LAVVULALPVAWGCCNAFEWLFFARTYNL TDEFEFFIGTYLYNFERGRGYSGSTLICS GOCKALVDTGASYSSKKLSTLAHHAMYPPPTPTT AGAYLPAWGCCNAFEWLFFARTYNL TDEFEFFIGTYLYNFERGRGYSGSTLICS VWTGAKDRCRRKSCRNPPDPVNGMYHVKG QFGSQKKYSCTKGSVBLIGSSATCIISGDTVIW DNETFICDRPCGLPTTTTNODFISTNERHY GSVVTYRCPGSGGRKVELQFSNCTSND DQVGWSGPAPQCIPNKCTPPNVENGILVSD NRSLSINEVVEVEGRCGSSATCIISGDTVIW DNETFICDRPCGLPTTTTNODFISTNERHY GSVVTYRCPGSGGRKVELQFSNCTSND DQVGWSGPAPQCIPNKCTPPNVENGILVSD NRSLSINEVVECHCGSSATCIISGDTVIW DNETFICDRPCGGGRKVELQFSNCTSND DQVGWSGPAPQCIPNKCTPPTVENGILVSD NRSLSINEVVECHCGSSATCIISGDTVIW DNETFICDRPCGGGRKVELQFSNCTSND DQVGWSGPAPQCIPNKCTPPNVENGILVSD NRSLSINEVVECHCCRSSATCIISGDTVIW DNETFICDRPCGGGRKVELVELGESTCCTSND DQVGWSGPAPQCIPNKCTPPNVENGILVSD NRSLSINEVVECHCGGSGGRCCHCALDGATCHQAPDHFI LEVTPFGKAVNYTCDDFDRGTSFDLIGGSTR CTSDPQGGCCCAGCHQCAPDHFI CTSDPQGGCCCAGCHCCHCAPQDAFD		Į.												
RKSYSWKQRLFIINFISFFSALAVYFRIIMYC EAGVYTTALLETYTVLTNMAPHMTAWWDF GNKELLTISQPEKRF  799 2149 A 6529 1 874 PFFPQRINTEHSGSVSLLALCDLGWCEDWS CCLVQGGGLVDWVQTNHGEDEAGGDTDSV DEARCKSQQEAQENLREDLCLESTAKKIL QIEGSERHEIETRIQAALDGEPLGGGQLTA VHLHPSKEQQGQEGGERQRGARTHHWRGW EKGRRVALRPSGKLRADPOPKRLGGFTPST ELPGLQPHAPTPHTAPATPTYSPADDTPNPV RWKCPLPVERFRTQLCERSTACPPKRPPL GLPODTTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK  800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKQYDSTIF TASKNGTYKPCFSNEPSTFTHKTVYPDQVG EITHLCFLVRDRVSAALTOMESACVSIHEALKS VIDYQTHFRLEAGGRSRAEDLNTRVAYWSV GEALLLLVVSIGQVELLSSPFSKKTTTTRVGS WIDYQTHFRLEAGGRSRAEDLNTRVAYWSV GEALLLLVSIGQVELLSSPFSKKTTTTRVGS WKRITLGNTTSSVLINYMDTQYTGEIGIGTP PQTFKVVPDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKINGTELTLRYSTGTVSGFL SQDIDTVGGITVTQMFGEVTEMPALPFMLAEP DOVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLOGQIVLGGDDPQHYE GNFHYNLIKTGVWQQMKGVSVGSSTLCE DGCLALVDTGASYISGSTSSIERLMEALGAKE KRLFDVYNKCNGGPTLPPTFLLGGKDTFLT SADYLFQESYSSKLSTLAHHAMYPPTPTTL ALGATFURKYTFETPGRNPHYGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPPVGPPAPGLFCAGFIL UAGATFURKYTFETPGRNPHGFALAR  804 MCLGRMGASSPRSPPVGPPAPGLFCAGFIL UAGATFURKYTFETPGRNPHGFALAR  805 MCLGRMGASSPRSPPVGPRAPGLFCAFFIL UAGATFURKYTFETPGRNPHGFALAR  806 MCLGRMGASSPRSPPVGPRAPGLIFCRNPHY UNDEFFICRPCGLPPTITINGDFISTNEHY GSVVTYRCPRGGGRKVECHANGVPHVIKG UGFGSQKYSCTKGYRLIGSSSATCINSGDTVIV DNETFICRPCGLPPTITINGDFISTNEHY GSVVTYRCPRGGGRKVECHANGVPHVIKG UGFGSQKYSCTKGYRLIGSSATCINSGDTVIV DNETFICRPCGLPPTITINGDFISTNEHY GSVVTYRCPRGGGRKVECHANGVPHVIKG UGFGSQKYSCTKGYRLIGSSATCINSGDTVIV DNETFICRPCGLPPTITINGDFISTNEHY GSVVTYRCPRGGGRKVECHANGPRVKCQA NRWEPELPSCSRVCQPPDVLHARRTQRDK DNSPAQPCCUFFRCOFFFVENNCTHFDRN DQVGIWSGPAPQCLIPNKCTFPNVENOLI VSD NRSLSLNEVVFCGFCCFSPTVINGGRTKPUCLGBSTR CTSDPQGNCVSSCAPCCGHCCSPTVINGGRTKFICLGBSTR CTSDPQGNCVSSCAPCCGHCCAPPDHT.		1												
BAGYYTRÁLEYTV/LIMMAFHMTAWWDF						j								
SOUND   STATE   STAT						1								
PFPPQRINFEHSGSVSLLALACDLGWCEDWS   CCLVQGGGGLVDVVQTIGEDEAGGDTDSV    CCLVQGGGGLVDVVQTIGEDEAGGDTDSV    DEARCKESQQEAQENLREDLCLESFAKDKIL     QIEGSEREHEETRIKQAALDGEPLGGGLTAV     HUFISKEQQGEGGGERGRARTHHWRGW     EKGRRVRLRPPSGKLRADQPVRKLGGPTPSYT     ELPOLOPHAPTHTAPATTYSPAPDTPNPPV     RWKCPLPVERTRQLCRERTRKACPPKPRPPL     GLPGDPTIGPVTHHAPPVSPTGASGQERRAEP     GAVSYAHASATK     GHYDVDCRLEDPDGKVLYKEMKKQVDSTFT     TASKNGTVKFCFSNESTHKTVYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVYYPDFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     VINTUTE     TASKNGTVKFCFSNESTHKTVTYPOFQVG     ETHLEFLVARDRVSALTQMESACVSHEALKS     TASKNGTVKFCFSNESTHKTVTYPOFQVG     ETHLEFLVARDRVSALTACVY     HKLFDASDSSSYKHOGTLILTRYSTGTVSGFL     SQDITVGGITVTQMFGEVTEMPALPFMLAEF     DGVVGMGFEQALGRVTFIFDNISGQVLKED     VFSTYYNRDSENSQLLGGQIVLGGSDPQHYE     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     DGCLALVDTGASTNSCTLLTRYSTGTVSGFL     SQDITVGGITVTQMFGEVTEMPALPFMLAF     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     DGCLALVDTGASTNSCTLLATART     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     DGCLALVDTGASTNSCTLATART     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     DGCLALVDTGASTNSCTLATART     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     DGCLALVDTGASTNSCTLATART     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSSVGSTLLCE     GNFHYNLIKTGVWQIQMKGVSVGSTLLCE     GNFHYNLIKTGVWGQVSSVLAGA     GNFHYNLIKTGV GNFHYNLI					•									
CCLVQGGGDLVDVVQTNIGEDEAGGDTDSV DEARCKESQQEAQENLREDLCLESPAKDKIL QIEGSEREHEITRTKQAALDGEFLGGGQLTA VHLHPSKEQQQGEGGERQRGARTHHWRGW EKGRRVRLRPPSGKLRADPVRKLGGPTPS/T ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVEPRTRQLCFRETKACPPKPRPPL GLPGDPTGVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK  800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGPGGAS EITFELPDNAKQCFVEDLAGGTKCTLEFQVTTG GHYDVDCALEDPDGKV,YEMKKQVDSFTF TASKNGTYKFCFSNEPSTFTHKTYYPDFQVG ETTHLCFLVR/DRVSALTQMESKAEDLNTKVATWSV GEALLLLVVSIGQVFLLKSFSDKRTTTRVGS  801 2151 A 6556 1 1319 TPCMECKGGGREPQTGEGSM DGWRNPRWGLLLLLWGSCTGGLPTDTTTF KRIFLKRMPSIRESLKERGLNTSVHALKS VIDYQTHERLRSAGGRSKAEDLNTKVATWSV GEALLLLVVSIGQVFLLKSFSDKRTTTRVGSF SQDITVGGTVTYQMFGEVTEMPALPFMLAEF DGWYCMGFEGALGRYPDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSSLYTACVY HKLPDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGTVTYQMFGEVTEMPALPFMLAEF DGVVGMGFEQAJGRYTPDNISQGVLKED VFSSYYNRDSENSQSLGGQTVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYBSSTSIEKLMEALGAKE KRLFDYVVKCNEGPTLPFTFLFLLGGGDTPLT SADVLFQESYSSKLSTLAHAMYIPPFTGFTL VLGGATPLRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGFPAPGLFFCCGGGL LAVVULALIPVAWGQCNAPEWLPFARPTNL TDEFEPFIGTYLNYECRPGYSGRPFSICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVKG IQFGSQIKYSCTKGYRLIGSSSATCISGDTVTW DNETPICDRECGLPPTTNGGDTVTSDD DQVGIWSGPAPQCIENKCTPPOVNGMVHVKG IQFGSQIKYSCTKGYRLIGSSSATCISGDTVTW GSVVTYRCNPGSGGRKVFELVGEPSIVCTSND DQVGIWSGPAPQCIPNKCTPPOVNGMVHVKG NFSPGGEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSQDFPMGGLINGRVLFPV NLQLGAKVDFVCDEFGQPVJKGRGRRYCCQA LNKWEPELPSCSRVQPPPDVLRGAASMRCTPQG DWSPAAPTCEVKSQDFPMGGLINGRVLFPV NLQLGAKVDFVCDEFGCPFVJKGRRVLFGLIGSSTI CTSDPQGNGWSSPAPTCEVGKGASSASVCVLAG MESLWNSSVPVCEQIFCSSPVIPNGRHTGKP LEYPFGKAVNYTCDPHPDRGTSFDLIGGSTI CTSDPQGNGWSSPAPTCEUGGSDLAGCQLAGCUNGCHPHCLTPGLIGSSTI CTSDPQGNGWSSPAPTCEUGHCDLGSSTI CTSDPQGNGWSSPAPTCEULGHCQAPDHFL	700	2140	<u> </u>	6520	1	874								
DEARCKESQQEAGENIAEDLCLESPAKDKIL  QIEGSEREHETRIKQALGEPILGGOQLTA VILHPSKEQQGQEGGERQRGARTHHWRGW EKGRVRLRPSGKLRADQFVRKILGGPTPS/T ELPCLQPHAPTHTHAPATPTYSPAPDTP)PPV RWKCPLPVERTRQLCREATRACQFYRSPPI GLOPDFIGOPTHHAPPYSPTGASGQERRAPE GAVSYAHASATK  800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLLALLLLVPGGGAS EITFELPDNAKQCFYEDLAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQVDSFIT TASKNGTYKFCSSNEWSTSTHIKTVYFDPQVG ETHLCFLVRDRVSALTQMESACVSHEALKS VIDYQTHEFRIREAQGRSEDLNTRVAYWSV GEALLLLVVSIGQVFLLKSFSDKRTTTTRVGS  801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLUWGSCTFGLPTDTTT KRIFLKRMFSIRESLKERGYDMARLGFWSQP MKRLTLGRITTSSVLLTNYDTQYYGEIGIGT PQTFKVVFDTGSSNVWYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTRYSTGTVSGFL SQDITTVGGTVTTQMFGEVTEMPALPPMLAEF DGVVGMGFIEQAIGRYTPIFDNISQGVLKED VFSFYYTRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSTLLCE DGCLALVDTGASYTSGSTSIEKLMEALGAKE KRLFDYVVKCNEGFTLPPTFLFLLGGKDTFL SADVLFQGSYSSKLISTAMAMYDPPTGFTL ALGATPIRKFYTEFDRGNNFHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSFEVGFPAFGLGFTCGGSL LAVVVLLALPVAWGQCNAPEWLPFARFINL TDEFEFPIGTYLNYECRPGYSGRFPSICLKNS VWTGAKDRGRKSSCRNPDPVNGMYHVKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTTTNOFTSTNEHPY GSVVTTKCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPCCIIPNKCTPPNVENGLIVSD NRSLFSLNENVEFTCQPGTVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLLAGASNRCTPQG DWSPAAPTCEVKSCDDFMGGLLNGRVLFPV NLQLGAKVDFVCDGFGFQLKGASSASYCVLAG MESLWNSSVPVCDQICFSPYUTNGRHTGCP LEVPFGKANNTCDPHDRGTSFTDLIGESTIR CTSDPQGNGWSSPAPCGLIGHCQAPDHFL	1	2177	1 ^	0525	) <b>'</b>	0/7								
OQIEGSERHETIKQAALDGEPLGGQQLTA V9LHPSKEQQQGGGERQARTHHWRGW EKGRRVRLRPPSGKLRADQPVRKLGOPTPS/T ELPGLQPHAPTPHTAPATPTYSPAPDTPNPPV RWKCPLPVEPRTRQLCRERTKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASQQERRAEP GAVSYAHASATK   OF SAME			Ì											
VILHPSKEQQQGEGĞERQRGARTHHWIĞUN	1													
BEGGRRYRLEPPSGKLRADOPYRKLGGPTPST   ELPGLOPHAPTPHTA/PATPTYSPAPDTPNPPV   RYKCPLPVEPRTRQLCRERTRKACPPKPRPPL   GLPGDPTGPVTHHAPPVSPTGASGQERRAEP   GAVSYAHASATK     800   2150   A   6544   2   662   SAQRWAAVGRWGCKLLALLLLVPGPGGAS   EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG   GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF   TASKNGTYKFCFSNEPSFTHKTYFDPQVG   ETHLCFLVRDRVSALTOMESACVSHEALKS   VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV   GEALLLLVISIQQVFLLKSFFSDKRTTTTRXQS     801   2151   A   6556   1   1319   TPCMECKIGEGLREPQNLSGSQREPQTEGSM   DGWRRMFRWGLLLLLWGSCTFGLFTDTTTT   KRIFKKRGGILREPNLSGSQREPQTEGSM   DGWRRMFRWGLLLLLWGSCTFGLFTDTTTT   KRIFKKRMPSIRESLKERGVDMARLGPEWSQP   MKRLTLGNTTSSVILLTNYMDTQYYGEIGIGTP   PQTFKVVPDTGSSNVWPSSKCSRLYTACVY   HKLFDASDSSYKINGTELTLRYSTGTVSGFL   SQDITVGGITVTQMFGEVTEMPALPFMLAEF   DGVVGMGFEQAIGRVTFIFDNISQGVLKED   VFSSTYTNADSENSQLAGGQIVLGGSDPQHYE   GNFHYTNLIKTGWQIQMKGVVSGSTLLCE   DGCLALVDTGASYJSSSIEKLMEALGAKE   KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT   SADYLFQESYSSKKLSTLAHAMYIPPTGFT   ALGATEMFKTYTETPDRGNPHGFALAR   ALGATEMFKTYTETPDRGNPHGFALAR   ALGATEMFKTYTETPDRGNPHGFALAR   ALGATEMFKTYTETPDRGNPHGFALAR   ALGATEMFKTYTETPDRGNPHGFALAR   GSVVTYRCNPGSGGRKVFELLVGERSYCTSDN   DQVGIWSGPAPQCIIRSSATCTISGDTVIW   DNETPICDRIPCGLPPTTTNODFISTNRENFHY   GSVVTYRCNPGSGGRKVFELLVGERSYCTSDN   DQVGIWSGPAPQCIIRNKCTPPNVENGILVSD   NRSLFSLNEVVEFRCQPFVMKGPRVKCQA   LNKWEPELPSCSRVCQPPDVLHAERTQRDK   DNESPGQEVTYSCEPGYDLRGAASMRCTPQQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGQLLNGRVLFPQ   DWSPAAPTCEVKSCDDFMGGLIGHCQAPDHTL   CTSDPQGMGVWSSSAPRCGIIGHCQAPDHFL   CTSDPQGMGVWSSSAPRCGIIGHCQAPDHFL   CTSDPQGMGVWSSSAPRCGIIGHCQAPDHFL   CTSDPQGMGVWSSSAPRCGIIGHCQAPDHFL   CTSDPQGMGVWSSSAPRCGIIGHCQAPDHFL   CTSDPQGMGVWSSSAPRCG														
BELPGLOPHAPTPHTA/PATTYSPAPDTPNPPY RWKCPLPVEPRTRQLCRETTR/ACCPPKPRPPL GLPCOPTGPVTHHAPPVSPTGASGQERRAEP GAVSYAHASATK   SAQRWAAVAGRWGCKILALLLLVPGPGGAS   EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLDPDGKVLYKEMKKQYDSFTF TASKNGTYKFCFSNENSTFHKTYYFDFQVG EVHLCFLVRDDRVSALTQMESACVSHHEALKS VDYQTHFRLREAQGRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS   SOI		ĺ												
RWKCPLPVERTRQLCRETTRKACPPKPRPPL GLPGDPTGPVTHHAPPVSPTGASGQERAAP GAVSYAHASATK  800 2150 A 6544 2 662 SAQRWAAVAGRWGCILLALLLVPGPGGAS EITFELPDNAXQCFVEDIAQGTKCTLEFQVTTG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKPCFSDREYSTFTHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSHIEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS  801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSGREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLQFEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVYPIFDNISQQVLKED VFSFYNRDSENSQSLGGQTVLGGSDPQHYE GNFHYNLIKTGVWQLQMKGVSVGSSTLLCE DGCLALVDTGASYSGSTSSIEKLMALGAKE KRLFDVVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFVGSYSSKKLSTLAHAMYIPPTGFTL VLGATFURKPYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPAGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFAPTNL TDEFEFFIGTYLNYCRPGYSGRFFSICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVKG QDGSVLYSCKGYGGSSSTICLCNS VWTGAKDRCRKSCRNPPDPVNGMVHVKG QDGSVLYSCKGYGGSSTCLINGSDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVSCTKGYLIGSSTATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGGSTICTISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLISGDTVIW DNETPICDRIPCGLPPTTINGDFISTNRENFHY GSVVTYRCAPGSGGRKVFLIGSGSATCLICR DNFSPGGEVFYSCEPGYDLGAAASMCTPQG DWSPAAPTCEVKSCDDPMGGLINGRVLPPV NLQLGAKVDFVCDGFGLKGSSASVCVLAG MESLWNSSVPVCEQIFCCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHDPRGTSFDLIGESTIR CTSDPQGNOWWSSPAPRCGLLCHCQAPDHFL CTSDPQGNOWWSSPAPRCGLLCHCQAPDHFL		l												
GLPGDPTGPVTHHAPPVSPTGASGQERRAEP   GAVSYAHASATK		1												
GAVSYAHASATK	1	ł	l	l ·	,									
800 2150 A 6544 2 662 SAQRWAAVAGRWGCRLIALLLLVPGPGGAS EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTVKFCPSNEFSTFITHTVYFDFQVG EITHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALLILLVVSIGQVFLLKSFFSDKRTTTTRVGS  801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRMPRWGLLLLWGSCTFGLPTDTTTF KRIFLKRMSIRESLKERGYDMARLGPEWSQP MKRITLGNTSSVLTTNYMDTQYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKINGTELTLRYSTGTVSGFL GOVVGMGFIEQAIGRVTPIFDNIISQGVLKED VPSFYYNRDSENSGSLGQGVL/GGSDPQHYE GNPHYNLLKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLPQESVSSKLSTLAHAMYPPPTGPTL VALGATTRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLFFARPTNL TDEFEPFIGTYLNYECRPGVSGRFFSICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYGSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNOFFISTNRENFHY GSVVTYRCNPGGSGRKVFELVGPSNVCCA LNK WEPELPSCXCOPPDVLHAERTQRDK DNSSPAPTCEVKSCDDFMQQLLNGRVLPPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQFFCGPSPVCIPRORLDRETCKP LEVFPFGKANNYTCDPHDVRGTFSPDLIGESTR														
EITTELPDNAKQCFYEDLAQGTKCTLEFQVITG GHYDVDCRLEDPDGKVLYKEMKKQVDSFITF TASKNGTVKFCPSNEPSTFITHKTVFPDFQVG EXTHLCFLVR/DRVSALTQMESACVSIHEALKS VIDYQTHEFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRWPRWGLLLLLWGSCTFGLPTDTTTF KRIFIKRMFSIESKLERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKINGTELTLRYSTGTVSGFL SQDITVGGTVTQMFGEVTEMPALPPMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENQSLGGQIVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYIGSTSSIEKLMEALGAKE KRLFDYVVKCHEGPTLPFTFLFLLGGKDTPLT SADYLFQESVSSKKLSTLAHIAMYIPPPTGFTL ALGAFFIRKFYTEFDRGNNPHGFALAR 802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARFINL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRSCRNPPDPVNGMYHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSTYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNSPGGGVFYSCEPGYDLRGAASSMRCTPQQ DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQFFCPSPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGGSTR	800	2150	A	6544	2	662								
GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF TASKNGTYKPCFSNEFSTFTHKTVYFDFQVG ENTHLCFLVKPORVSALTQMESACVSHEALKS VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS  801 2151 A 6556 1 1319 TPCMECIKGGGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNISGGVLKED VFSFYYNRDSSNQSLGGQIVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYNSGSTSSIEKLMEALGAKE KRLPDYVVKCNEGPTLPPTF-IFLLGKDTPLT SADYLFQESYSSKKLSTLAHAMYIPPPTGPTL VALGATFURKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAFGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPSIICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPFTTINGDFISTNRENFHY GSVVTYRCNPGSGGRKVELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDGGFQLKGSSASYCVLAG MESLWNSSVPVCEQIECPSPPVIPNGRHTIGKP LEVFPFGKAVNYTCDPHIPDRGTSFDLIGESTIR CTSDPQGMGVWSSPAPRCGILGHCQAPDHFL	""		l ''	***	-	002								
TASKNGTYKFCFSNEFSTFHKTVYFDFQVG ETHLCFLVR/DRVSALTQMESACVSHEALKS VIDYQTHFRIREAQGRSRAEDLNTRVAYWSV GEALILLVSIGQVFLLKSFFSDKRTTTTRVGS 801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSGREPQTEGSM DGWRRMPRGLILLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILINYMDTQYYGEIGIGTP PQTFK-VYFDTGSSNV-WYPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITTVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYTNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADVLFQESYSSKKLSTLAHAMYIPPTGFTL VALGATFIRKYTFEDTGGNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLFFCCGGSL LAVVVLALPVAWGQCNAPEWLPFARPTNL TDEFEFFIGTYLNYCECRGYSGRPFSILCKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETFICDRIPGJFTNTMGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVYFFCQGPFWMKGPRRVKCQA LNKWEPELPSCSRVCQPPPVLHAERTQRK DNFSPGGEVFYSCEFGYDLRGAASMRCTPQQ DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDGFFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPVIPPIGRHTIGKP LEVFPFGKAVNYTCDPHDRGTSFFDLIGESTTR CTSDPQGMGVWSSPAPRCGILGHCQAPDHFL														
801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM GEALILLVSIGQVFLLKSFFSDKRTITTRVGS REFINENCE AND CONTROL OF THE REPORT OF THE REPORT OF THE REPORT OF THE REFINENCE AND CONTROL OF THE RE														
801 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM  BOI 2151 A 6556 1 1319 TPCMECIKGEGLREPQNLSGSQREPQTEGSM  DGWRRMPRWGLLLLWGSCTFGLPTDTTTF  KRIFLKRMFSIRESIKERGVDMARLGFEWSQP  MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP  PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY  HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL  SQDITTVGGTVTQMFGGVTEMPALPFMLAEF  DGVVGMGFIEQAIGRVTPIFDNISQGVLKED  VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE  GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE  DGCLALVDTGASYISGSTSISEKLMEALGAKE  KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT  SADYLFQESYSSKKLSTLAHAMYIPPPTGPTL  ALGATFIRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL  LAVVVLLALPVAWGQCNAPEWLPFARPTNL  TDEFEFFIGTYLNYECRPGYSGRPFSIICLKNS  VWTGAKDRCRKSCRNPDPVNGMVHVIKG  IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW  DNETFICIDRIPCGLPPTITNGDFISTNENFHY  GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND  DQVGIWSGPAPQCIPNKCTPPNVENGILVSD  NRSLFSLNEVVEFRCQPGFVMKGPRVKCOA  LNKWEPELPSCSRVCQPPPDVLHAERTQRDK  DMSPAAPTCEVKSCDDFMGQLLNGRVLFPX  NLQLGAKVDFVCDEGFQLKGSSASYCVLAG  MESLWNSSVPVCEGIFCPSPPVIPNGRHTGKP  LEVFPFGKAVNYTCDPHDRGTSFDLIGESTIR  CTSDPQGNGWWSSPAPRCGILGHCQAPDHFL							l							
801 2151 A 6556 1 1319 TPCMECIKGEGLEPQNLSGSQREPQTEGSM DWRRMPWGLLLLLWGSCTFGLPTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVILINTYMDTQYYGEIGIGTP PQTFK VVFDTGSSNV WVPSSKCSRLYTACVY HKLFDASDSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFLEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGOIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDVVKKCNEGPTLPPTFFLLGGKDTPLT SADYLFQESYSSKKLSTLAHAMYIPPTGPTL VALGATFURKFYTEFDRGNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPSPEVGPPAPGLFFCCGGSL LAVVULLALPVAWGQCNAPEWLPFARPTNL TDEFFFFIGTYLNYECRPGYSGRPFSICLKNS VWTGAKDRCRKSCRNPPDPVNGMVHVIKG IQFGSQLKYCKGYRLIGSSATCIISGDTVIW DNETPICDRIPCGLPPTTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVQPPPDVLHAERTQRDK DNFSPQAEVFTSCPGFYDLRGAASMRCTTQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESL WNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKANNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL		1												
801 2151 A 6556 1 1319 TPCMECIKGEGIREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLFTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQP MKRLTLGNTTSSVLTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGTTVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLPQESYSSKLSTLAHIAMYIPPPTGPTL ALGATPIRKFYTEFDRGNNPHGFALAR ROLL VDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTGRGNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEPFIGTTYNYECRPGYSGRPFSICLKNS VWTGAKDRCRRESCRIPPDPVMGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRPCGLPPTITNOFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNK WEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCQIECPSSPVIPNGRNTGKP LEVFPFGKAVNTYCDPIPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPCGILGHCQAPDHFL		ľ	ì			• .								
DGWRRMPRWGLLLLLWGSCTFGLFTDTTTF KRIFLKRMPSIRESLKERGVDMARLGPEWSQD MKRLTLGNTTSSVLTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGGIVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRYTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRFFSILCLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPFNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPQGEVFYSCEPGYDLRGAASMRCTPQG DWSFAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVUPNGRHTIGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL	801	2151	A	6556	1	1319								
MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP PQTFKVVFDTGSSNVWPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAFGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVPFFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDITVGMFGEVTEMPALPFMLAEF DGVVGMGFEQAIGRVTPIFDNISSQVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSILCLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVPFFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL							KRIFLKRMPSIRESLKERGVDMARLGPEWSQP							
HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL SQDIITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYNLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITINGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQILNGRVLFPV NLQLGAKVDFVCDGEFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL					-		MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP							
SQDITVGGITVTQMFGEVTEMPALPFMLAEF DGVVGMGFEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTINRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL	,				-		PQTFKVVFDTGSSNVWVPSSKCSRLYTACVY							
DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFFFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL							HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL							
VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL ALGATFURKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDGFFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL		ļ												
GNFHYINLIKTGVWQIQMKGVSVGSSTILCE DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATFIRKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTTNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL	1 .		ł				DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED							
DGCLALVDTGASYISGSTSSIEKLMEALGAKE KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\RKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
KRLFDYVKCNEGPTLPTFLFLLGGKDTPLT SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATF\RKFYTEFDRGNNPHGFALAR  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEW\LPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL			,				• •							
SADYLFQESYSSKKLSTLAIHAMYIPPTGPTL VALGATF\(\text{IRKFYTEFDRGNNPHGFALAR}\)  802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL				,										
802 2152 A 6567 13 6147 MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
LAVVVLLALPVAWGQCNAPEWLPFARPTNL TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
TDEFEFPIGTYLNYECRPGYSGRPFSIICLKNS VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL	802	2152	Α	6567	13	6147								
VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
DNETPICDRIPCGLPPTITNGDFISTNRENFHY GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL					i									
DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDGEFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
LNKWEPELPSCSRVCQPPPDVLHAERTQRDK DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL						1								
DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL	·													
DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL					j									
NLQLGAKVDFVCDEGFQLKGSSASYCVLAG MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL														
LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL				•	. [									
CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL					.									
				6		*								
							FAKLKTQTNASDFPIGTSLKYECRPEYYGRPF							

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	- build	l	914		acid residue	
uence	1	1	714	ng to first		Q=Glutamine, R=Arginine, S=Serine,
•	i	l	ļ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		Ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion, \=possible
1	ł	l	ł	sequence		nucleotide insertion
	1					SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG
1	1	į	1			MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI -
1	ľ		1		1	LSGNAAHWSTKPPICORIPCGLPPTIANGDFIS
			i			
			1			TNRENFHYGSVVTYRCNPGSGGRKVFELVGE
1	1		1		ĺ	PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV
	l l	1	1			ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP
1	ì	İ	]			RRVKCQALNKWEPELPSCSRVCQPPPDVLHA
	İ				•	ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS
		1				MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN
	ł		i	i		GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS
1		1				YCVLAGMESLWNSSVPVCEQIFCPSPPVIPNG
i		I		}		7
1	1	ļ	]		]	RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD
1		1				LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC
1		1			•	QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP
1		l				EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP
1		·				PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG
i		I	1	1		HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI
1		ŀ				ANGDFISTNRENFHYGSVVTYRCNLGSRGRK
i		ľ	1			VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN
			[			KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP
		Ì				GFVMKGPRRVKCQALNKWEPELPSCSRVCQ
						PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY
						DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF
						LGQLPHGRVLFPLNLQLGAKVSFVCDEGFRL
			l :			KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN
1		ł				PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR
ĺ						
1	1		i			GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE
1	i					LSVRAGHCKTPEQFPFASPTIPINDFEFPVGTS
	1					LNYECRPGYFGKMFSISCLENLVWSSVEDNC
1						RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC
1						NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII
1	[					SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH
	1					TGPDGEQLFELVGERSIYCTSKDDQVGVWSS
			1			
1						PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI
1						IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH
1					ļ	CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY
1						SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV
)	]					KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC
Į	]		{			DEGFRLKGRSASHCVLAGMKALWNSSVPVC
1						EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT
İ	]				.	CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS
1						
1		•	!		l	SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL
1	[			-		YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS
1	]			i	ľ	QLDHYCKEVNCSFPLFMNGISKELEMKKVYH
1	1 !					YGDYVTLKCEDGYTLEGSPWSQCQADDRWD
						PPLAKCTSRTHDALIVGTLSGTIFFILLIIFLSWI
1	į l					ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP
						RTLOTNEENSRVLP
803	2153	Α	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL
1 000	2133	^	3374	-	,2,3,5	
		ļ				LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY
	'					PWSWA\RVGPAVELALAQVKARPDLLPGWT
			<u> </u>			VRTVLGSSENALGVCSDTAAPLAAVDLKWE
1					]	HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL
1						TAGAPALGFGVKDEYALTTRAGPSYAKLGDF
						VAALHRRLGWERQALMLYAYRPGDEEHCFF
				[	,	LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT
1			]			
		[		1		RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA
1						GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW
			1	ļ		ERGDGQDVSARQAFQAAKIITYKDPDNPEYL
1 3						FET VOLVIE AND COMPANDA CONTRACTOR
		- 1	1	I	I	EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	bod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	scq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	[		[	peptide	1 4	/=possible nucleotide deletion, \=possible
	<b>,</b>			sequence		nucleotide insertion
					i -	DGLLLYIQAVTETLAHGGTVTDGENTTQRMW
						NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
						NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
						YPPPDIPKCGFDNEDPACNODHLSTLEVLALV
						GSLSLLGILIVSFFIYRKMQLEKELASELWRVR
					•	WEDVEPSSLERHLRSAGSRLTLSGRGSNYGSL
1	ļ	i				LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
						IELTRKVLFELKHMRDVQNEHLTRFVGACTD
1	1					PPNICILTEYCPRGSLQDILENESITLDWMFRY
i l					9	SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
						DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
						LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
<b>i</b>						IALRSGVFHVEGLDLSPKEHERVTRGEQPPFR
						PSLALQSHLEELGLLMQRCWAEDPQERPPFQ
1						QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
1				İ		EELVEERTQAYLEEKRKAEALLYQILPHSVAE
						QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE
1 1						STPMQVVTLLNDLYTCFDAVIDNFDVYKVET
						IGDAYMVVSGLPVRNGRLHACEVARMALAL
]						LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV VGLKMPRYCLFGDTVNTASRMESNGEAL\KI
1						HLSS\ETKAVL\EEFGGFELELRGDVEMKGKG
1						KYRTYWLLGERGSSTRG
804	2154	A	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
						MSERVSGLAGSIYREFERLIVRYDEEVVKELIP
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE
1	ŀ		1			QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
1 1						DLQTRVESLESQTRQLELKAKNYADQISILEE
1 1						REAELKKEYNALHQRHTEMIHNYMEHLERT
1 1						KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
			İ	ļ		AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
						SHTSLKDELSDVSQGGSKATTPASTANSDVA
		í		i	•	TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
]		Į	]		ļ	QVAQETRNVSTGSAENEEKSEVQAIIESTPEL DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
1	1					FEELSSAGSGLIGDVDEGADLLGMGREVENLI
	ļ					LENTQLLETKNALNIVKNDLIAKVDELTCEK
1 1	l	ŀ				DVLQGELEAVKQAKLKLEEKNRELEEELRKA
1		. [				RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
		į				MARVLMERNOYKERLMELOEAVRWTEMIR
1		l	}	į		ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK
		į				KPEPPVNLKYNAPTSHVTPSVKKRSSTLSQLP
						GDKSKAFDFLSEETEASLASRREQKREQYRQ
1	i	ł		ı	}	VKAHVQKEDGRVQAFGWSLPQKYKQVTNG
		ļ		[		QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
	ŀ	1	!	1		VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
1 1	{		- 1	- 1		TEGSKQRSASQSSLDKLDQELKEQQKELKNQ
	İ	l				EELSSLVWICTSTHSATKVLIIDAVQPGNILDS
		l				FTVCNSHVLCIASVPGARETDYPAGEDLSESG
	ĺ	1	ſ	. 1	1	QVDKASLCGSMTSNSSAETDSLLGGITVVGC
	1	l		l		SAEGVTGAATSPSTNGASPVMDKPPEMEAEN SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS
]	İ	1		1		QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND
[	ſ	- 1		(	1	SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
	l	ł		ĺ	İ	MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
		}	1	ı	ļ	SILSIVHVKGIVLVALADGTLAIFHRGVDGQW
	1	1	9,0		ł	DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
	ľ	ł	1	1	- 1	YRNKIYVVQPKAMKIEKSFDAHPRKESQVRQ
			1			LAWVGDGVWVSIRLDSTLRLYHAHTYQHLQ
L				·		DVDIEPYVSKMLGTGKLGFSFVRITALMVSC

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq- uence	uence		09/496 914	correspondi ng to first amino acid residue of	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
				Soques		NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
						ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
805	2155	A	6605	469	2602	ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP   FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
						SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
						SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ
						DSGLYACVIRNSTYCMKVSISLTVGENDTGL CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
						REPEILWYKECRTKTWRPSIVFKRDTLLIREV REDDIGNYTCELKYGGFVVRRTTELTVTAPL
						TDKPPKLLYPMESKLTIQETQLGDSANLTCRA FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
	-					SDI\KILKEHLGEQEVSISLIVDSVEEGDLGNYS
						CYVENGNGRRHASVLLHKRELMYTVELAGG LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA
						EELDGDNKDYDAYLSYTKVDPDQWNQETGE EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT
						YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF ELETRLRNMLVTGEIKVILIECSELRGIMNYQE
		,				VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
		*			•	LQYEMPFKRIEPITHEQALDVSEQGPFGELQT VSAISMAAATSTALATAHPDLRSTFHNTYHS
						QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
806	2156	Α	6614	3	1584	ILPLLPRETSISSVIW NSARGGVGVRGARAMATVQEKAAALNLSAL
500	2150	^	0014	3	1507	HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT
						QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL IQNKYFGDVDIPRAKVVRVCQALMDYKVFE
						AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQD SQLGKENKLYSPARYADALFKSSDIRSASLED
						LWENLSLKPANSPHVNISTTLSPQVINEVWQE ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
						RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA
						AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
			-			DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT L\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRI
						DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
807	2157	A	6615	4198	2094	KEKKKLLGQFYKCHPDIFIEHFGD FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
	ļ					PQILLLALLTLGLAAQHQDKVPCKM/VKML CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD
						LSGNOLRSILASPLGFYTALRHLDLSTNEISFL OPGAFOALTHLEHISLAHNRLAMATALSAG
						GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS
		. v				LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD
						FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR
						LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
		•				GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE
i.		-				TLELGARALG\SLRTLLLQGNALRDLPPYTFA NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\
						AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	İ			peptide	sequence	/=possible nucleotide deletion, \=possible
	İ	Ì		sequence		nucleotide insertion
				5545555	<u> </u>	LDLSSNPGLEVATGALGGLEASLEVLALQGN
						GLMVLQVDLPCFICLKRLNLAENRLSHLPAW
	1		!			TQAVSLEVLDLRNNSFSLLPGSAMGGLETSLR
		1	ŀ			RLYLQGNPLSCCGNGWLAAQLHQGRVDVDA
	ł		ŀ		l	TQDLICRFSSQEEVSLSHVRPEDCEKGGLKNI
			1			NLIILTFILVSAILLTTLAACCCVRRQKFNQQ
						YKA
808	2158	A	6619	153	1852	FKALSQYIYTNTHLEREAAFEVAILLRRMEEG
						ARHRNNTEKKHPGGGESDASPEAGSGGGGV
						ALKKEIGLVSACGIIVGNIIGSGIFVSPKGVLEN
1	1	1	ł		1	AGSVGLALIVWIVTGFITVVGALCYAELGVNI
						PKSGGDYFYVKDIFGGLAGFLRLWIAVLVIYP
						TNQAVIALTFSNYVLQPLFPTCFPPESGLRLLA AICLLLLTWVNCSSVRWATRVQDIFTAGKLL
						ALALJIMGIVQICKGEYFWLEPKNAFENFOEP
[						DIGLVALAFLQGSFAYGGWNFLNY\VTEELV
1						DP\YKNL\PRAIFISIP\LVTFVYVFANV/ALYVT
					ĺ	AMSPQEL\LAS\NAVAVTFGEKLLGVMAWIM
						PISVALSTFGGVNGSLFTSSRLFFAGAREGHLP
						SVLAMIHVKRCTPIPALLFTCISTLLMLVTSD
·					•	MYTLINYVGFINYLFYGVTVAGQIVLRWKKP
						DIPRPIKINLLFPITYLLFWAFLLVFSLWSEPVV
						CGIGLAIMLTGVPVYFLGVYWQHKPKCFSDFI
						ELLTLVSQKMCVVVYPEVERGSGTEEANED
809	2159	_	6621	1041	223	MEEQQQPMYQPTPTKDKDVAGQPQP
809	2139	A	0021	1041	223	QDSRKMLPSTSVNSLVQGNGVLNSRDAARH
						TAGAKRYKYLRRLFRFRQMDFEFAAWQMLY LFTSPQRVYRNFHYRKQTKDQWARDDPAFL
						VLLSIWLCVSTIGFGFVLDMGFFETIKLLLWV
1 1						VLIDCVGVGLLIATLMWFISNKYLVKRQSRD
						YDVEWGYAFDVHLNAFYPLLVILHFIQLFFIN
						HVILTDTFIGYLVGNTLWLVAVGYYIYVTFL
] ]						GYSVGLLFFS\ALPFLKNTVILLYPFAPLILLYG
						LSLALGWNFTHTLCSFYKYRVK
810	2160	A	6623	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAA
						QQVAEDKFVFDLPDYESINHVVVFMLGTIPFP
	· [					EGMGGSVYFSYPDSNGMPVWQLLGFVTNGK
					-	PSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQ
						IGISVELLDSMAQQTPVGNAAVSSVDSFTQFT
						QKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMF
				ł		IPANVVLKWYENFQRRTSTEPSLLENIIWIKIN F
811	2161	Ā	6627	18	3367	LEGSLNTERAKYYLTITMPHFTVTKVEDPEEG
"	2101	^	5027	.0	3307	AAASISQEPSLADIKARIQDSDEPDLSQNSITG
	1					EHSQLLDDGHKKARNAYLNNSNYEEGDEYF
1 1		- 1		ł	l	DKNLALFEEEMDTRPKVSSLLNRMANYTNLT
		ŀ		l		QGAKEHEEAENITEGKKKPTKTPOMGTFMG
	İ			]		VYLPCLQNIFGVILFLRLTWVVGTAGVLQAF
				l		AIVLICCCCTMLTAISMSAIATNGVVPAGGSY
	-	[		[	[	FMISRALGPEFGGAVGLCFYLGTTFAAAMYIL
		ļ		]		GAIEIFLVYTVPRAAIFHSDDALKESAAMLNN
]		l				MRVYGTAFLVLMVLVVFIGVRYVNKFASLFL
]	Į		}	ļ	ļ	ACVIVSILAIYAGAIKSSFAPPHFPVCMLGNRT
		ļ		ļ		LSSRHIDVCSKTKEINNMTVPSKLWGFFCNSS
		ļ				QFFNATCDEYFVHNNVTSIQGIPGLASGIITEN
( i		- 1	ĺ	ſ		LWSNYLPKGEIEKPSAKSSDVLGSLNHEYVL
		ļ		ŀ		VDITTSFTLLVGIFFPSVTGIMAGSNRSGDLKD AQKSIPIGTILAILTTSFVYLSNVVLFGACIEGV
		i	ļ		}	VLRDKFGDAVKGNLVVGTLSWPSPWVIVIGS
Į l	J	J	}	ļ	}	FFSTCGAGLQSLTGAPRLLQAIAKDNIIPFLRV
					1	TEAL INTERNATION TO SECOND LINE LEVA

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		0 <del>9</del> /496	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
dence	Í	l	314	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide	Jaquan	/=possible nucleotide deletion, \=possible
]				sequence		nucleotide insertion
						FGHSKANGEPTWALLLTAALAELGILLASLDL
	1	l		ì		VAPILSMFFLMCYLFVNLACALQTLLRTPNW
	<b> </b>	ĺ			ĺ	RPRFRYYHWALSFMGMSICLALMFISSWYYA
						IVAMVIAGMIYKYIEYQGAEKEWGDGIRGLS
	İ	l			i	LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV
	]	ļ	]			GNFLENYGEALAAEQTIKHLMEAEKVKGFCO
						LVVAAKLREGISHLIQSCGLGGMKHNTVVM
	į	l				GWPNGWRQSEDARAWKTFIGTVRVTTAAHL
		Į.				ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG
1	]		1			GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE
						DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS
1		ł		}		DISAYTYERTLMMEQRSQMLRHMRLSKTER
		]				DREAQLVKDRNSMLRLTSIGSDEDEETETYQ EKVHMTWTKDKYMASRGQKAKSMEGFODL
						LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA
		1	j			KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL
			[			ERVLLVRGGGSEVITTYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG
1			1			DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP
					,	MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ
						VAGADFESEDEGEEFDDWEDDYDYPEEEQLS
						GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELF\SLMVVNRLTEELG
1						CDEIIDRE
813	2163	Α	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC
						WQYRQLSALHRAPRPTRPDKARRLGYKAKQ
						GY/VYIYIGFVFAVIYRIRVRRGGRKRPVPKG
						ATYGKPVHHGVNQLKFARSLQSVAEERAGR
}						HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS
						RGLGKGHKFHHTIGGSRRAAWRRNTLQLH
] ]						RYR
814	2164	A	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR
						DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT
						AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN
					İ	PLTKESIRQKEMESKRLRLLQEEDRRKKIARM
1 .						GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMEDICKVOIDSMDDSALASDDENI II ADTNIS
						CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF
						MHENLYFINRKVINSVCWASLNHLDSHILLC
						LMGLAETPGCATLLPASLFVNSHPAGIDRPG\
					l	MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR
	1					RVLLTNVVTGHRQSFGTNSDVLAQQFALMA
						PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF
•		1		1		HDSAVTSVRILQDEQYLMASDMAGKIKLWD
			ľ			LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA
	ļ		1			DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY
			•			SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG
	i					PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE
	ł	- 1	ł	į	ŀ	RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT
			Ì			PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF
			- 1			QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA
j j		J	1			LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR
	i		- 1		[	VSVDWGKCMNPFRNMVLEILDVSGNGWTV
	İ	i	1		j	DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN
_						IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  RVFETLKDLKVLNLAYNKINKIADEAFYGLD NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR LENLDILYFLLRVPHLQILILNQNRFSSCSGDQ TPSENPSLEQLFLGENMLQLAWETELCWDVF EGLSHLQVLYLNHNYLNSLPPGVFSHLTALR GLSLNSNRLTVLSHNDLPANLEILDISRNQLL APNPDVFVSLSVLDITHNKFICECELSTFINWL NHTNVTIAGPPADIYCVYPDSLSGVSLFSLSTE GCDEEEVLKSLKFSLFTVCTVTLTLFLMTILTV TKFRGFCFICYKTAQRLVFKDHPQGTEPDMY KYDAYLCFEERDFVPGENRPANIQDAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKEKK
816	2166	A	6646	1	3811	KDNNIPLQTVATIS  RDRAGVRPAGKQHAAAAFYDVGGDRPWDS
						GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQP KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGKVCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDDD
817	2167	<u>A</u>	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQLLGLIQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR
818	2168	A	6660	357	1890	KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED  APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYTDNL WDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPGHLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCEELLLPLENPCQGRA RLLLTGLHACG/DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSPGFHAELLPIPSPELSPRNLVLVATKMPLG QALSVLETEDS
819	2169		6661	65	2686	SGSGHCLAEAASMGPWGWKLRWTVALLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRQFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGTTLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARDMRSCLTEGEAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAGNGNYEKKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFPKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EMTNLKDIGLYNLRNITRG\AIRIEKNADLCYL STVDWSLILDAVSNNYIVGNKPPKECGDLCP GTMEEKPMCEKTTINNEYNYRCWTTNRCQK MCPSTCGKRACTENNECCHPECLGSCSAPDN DTACVACRHYYYAGVCVPACPPNTYRFEGW RCVDRDFCANILSAESSDSEGFVHDGECMQE CPSGFIRNGSQSMYCIPCEGPCPKVCEFEKKT KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA SELENFMGLIEVVTGYVKIRHSHALVSLSFLK NLRLILGEEQLEGNYSFYVLDNQNLQQLWD WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE VTGTKGRQSKGDINTRNNGERASCESDVLHF TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK EAPFKNVTEYDGQDACGSNSWNMVDVDLPP NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
						VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE AEYRKVFENFLHNSIFVPRPERKRRDVMQVA NITMSSRSRNITAADTYNITDPEELETEYPFF ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE KLGCSASNFVFARTMPAEGADDIPGPVTWEP RPENSIFLKWPEPENPNGLILMYEIKYGSQVE DQRECVSRQEYRKYGGAKLNRLNPGNYTARI QATSLSGNGSWTDPVFFYVQAKRYENFIHLII ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN GVLYASVNPEYFSAADVYVPDEWEVAREKIT MSRELGQGSFGMVYEGVAKGVVKDEPETRV AIKTVNEAASMRERIEFLNEASVMKEFNCHH VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR SLRPEMENNPVLAPPSLSKMIQMAGEIADGM AYLNANKFVHRDLAARNCMVAEDFTVKIGD FGMTRDIYETDYYRKGGKGLLPVRWMSPESL KDGVFTTYSDVWSPGVVLWEIATLAEQPYQ GLSNEQVLRFVNMEGGLLDKDNCPDMLFEL MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE VSFYYSEENKLPEPEELDLEPENMESVPLDPS ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF DERQPYAHMNGGRKNERALPLPQSSTC
821	2171	A	6691	106	825	GRVLFRGCGVGHKGQVLMGTFILAQDWLSE SNHVFCVSSMLRLQKRLASSVLRCGKKKVW LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT VHSRARCRKNTLARRKGRHMGIGKRGTAN ARMPEKVTWMRRMRILRRLLRRYRES/KRYR ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH IHKLKADKARKKLLADQAEARRSKTKEARK RREERLQAKKEEIIKTLSKEEETKK
822	2172	A .	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI SGAKSSSS\RFTDSKRYEK\DFQ\SCFGLHETR\ SGDI\CNA\CVLL\LKRWKKLPAGSKK\NWNH VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\ TYWKRQKICCG\\YKGRFGEVLIDTHLFKPCC SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	Ą	6727	3	4063	PYLATLQLDSSLLIPKYQTPPAAAQGQATPG NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS VGGQNPSTGGISADRTQGNIGCGGDTDPGQS

C 655	1 000 000	157:	CEC	I n 3** - 3	18-11-	1 A 2 A 2 A 2 A 2 A 2 A 2 A 2 A 2 A 2 A
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i	l	914	ng to first	acid residue ·	Q=Glutamine, R=Arginine, S=Serine,
l		ĺ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i	Į	ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į		1		peptide		/=possible nucleotide deletion, \=possible
į	ļ	l	i	sequence		nucleotide insertion
			<del> </del>	sequence	<del></del>	SSQPSQDGQESNVPSVGSLADPDYLNTPOMN
1				,	1	
					1	TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP
		· ·	ł .			RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP
1	ĺ	ĺ	ĺ			ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL
1		1				SDSVMNIFKDRNFDSCCICACNMNIKGADVG
		ļ				LYIPDSSNEDQYRCTCGFSAIMNRKLGYNSGL
	l	İ		·	i	FLEDELDIFGKNSDIGQAAERRLM\MCQSTFL
					1	PQVEGTKKPQEPPISLLLLLQNQHTQPFASLN
1	i	1	1			FLDYISSNNRQTLPCVSWSYDRVQADNNDY
1	ļ	1				WTECFNALEQGRQYVDNPTGGKVDEALVRS
1		ł				ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP
ļ					i	
		l				FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH
	İ	1	1			KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT
	1	Ì				ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN
		l	•			EALLEGAKTFFRDLSAVYEMCRLGQHKPICK
1		1			}	VLRDGIMRVGKTVAQKLTDELVSEWFNQPW
İ		l				SGEENDNHSRLKLYAQVCRHHLAPYLATLQL
		1				DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN
1	ŀ	ł				GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV
1	l	Ì				PPVSSSASAPGISOISTTSSSGFSGSVGGONPST
]	!	l				GGISADRTOGNIGCGGDTDPGOSSSOPSODG
l		l				OESVTERERIGIPTEPDSADSHAHPPAVVIYM
į		ĺ				7
1						VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD
ļ		i				NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY
1	l	ļ	l i			IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP
						AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT
· ·						ELGETFGEASQKYNVLFVGYCLSHDQRWLL
		i				ASCTDLHGELLETCVVNIALPNRSRRSKVSAR
						KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR
1	l					LGHGELKDWSILLGECSLQTISKKLKDVCRM
İ	1	İ		7		CGISAADSPSILSACLVAMEPQGSFVVMPDAV
	}		l			TMGSVFGRSTALNMQSSQLNTPQDASCTHIL
			1			VFPTSSTIOVAPANYPNEDGFSPNNDDMFVDL
1	1		ľ			`
						PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP
						SGIGVGSHFQHSRSQGERLLSREAPEELKQQP
1	1					LALGYFVSTAKAENLPQWFWSSCPQAQNQC
						PLFLKASLHHHISVAQTDELLPARNSQRVPHP
1			1	1	}	LDSKTTSDVLRFVLEQYNALSWLTCNPATQD
						RTSCLPVHFVVLTQLYNAIMNIL
824	2174	Α	6732	2440	365	VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
						RLLPPSSAAAFSSHRHNLLCSRRRGGGGGG
1		l '				GGGGGTIKRPGITGPTAATSPSGEPGNAASAP
1	}			1	1	LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC
1						ASLVFGRLQHRGGDRKRGLLGRSSGDAASD
1						OPFRCRSGSTAGRLVKOMDFTEAYADTCSTV
						,
						GLAAREGNVKVLRKLLKKGRSVDVADNRG
						WMPIHEAAYHNSVECLQMLINADSSENYIKM
						KTFEGFCALHLAASQGHWKIVQILLEAGADP
1						NATTLEETTPLFLAVENGQIDVLRLLLQHGAN
	l			İ		VNGSHSMCGWNSLHQASFQENAEIIKLLLRK
						GANKECQDDFGITPLFVAAQYG\KLESL\SILIS
						SG\ANVNCOALDKATPLFIAAOEGHTKCVELL.
						SG\ANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWOLPIHAAAOMGHTKI
0						LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI
0				,		LSSGADPDLYCNEDSWQLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE
						LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ
						LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK
				•		LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA
				·		LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT
						LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA
		<b>3</b> 5				LSSGADPDLYCNEDSWOLPIHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQINELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalamine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG
					·	TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRHNIILSAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARA\TDYLQ ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	2176	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFEGEKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFFGHQKLCQFCIS SMVQQGIQIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFFRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFITVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRLTL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSHASLKLYI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGGGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1100	in in	nucleotide		
eotide		ļ			location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	1	USSN	location	corresponding	I-Isoleucine, K-Lysine, L-Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ŀ		Į i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ŀ	ł	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ	l		peptide	-	/=possible nucleotide deletion, \=possible
1		l		sequence		nucleotide insertion
	f					VGKGIMGVFTKPIGGAAELVSQTGYGILHGA
1				ì		GLSQLPKQRHQPSD\VHADQAPNSHVKYVW
1	J .		]			KMLQSLGRPEVHMALDVVLVRGSGQEHEGC
						LI I TOPUL EUROPEDTOOOA ERVITORIO A OR
1			1			LLLTSEVLFVVSVSEDTQQQAFPVTEIDCAQD
	1			i		SKQNNLLTVQLKQPRVACDVEVDGVRERLSE
	1					QQYNRLVDYITKTSCHLAPSCSSMQIPCPVVA
1	(		i i			AEPPPSTVKTYHYLVDPHFAQVFLSKFTMVK
						NKALRKGFP
827	2177	A	6748	2	1662	FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK
			[			ASGVSPTLWRPQAAATGLEMPSSGRALLDSP
			]			LDSGSLTSLDSSVFCSEGEGEPLALGDCFTVN
1						VGGSRFVLSQQALSCFPHTRLGKLAVVVASY
I						
1				<u>'</u>		RRPGALAAVPSPLELCDDANPVDNEYFFDRS
			[			SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI
1			1			QYWGIDELSIDSCCRDRYFRRKELSETLDFKK
						DTEDQESQHESEQDFSQGPCPTVRQKLWNIL
]			]	}		EKPGSSTAARIFGVISHFVGVSHNMALMSAEL
<b>.</b>			<b>!</b>			SWLDLQLLEILEYVCISWFTGEFVLRFLCVRD
			<b>!</b>			RCRFLRKVPNIIDLLAILPFYITLLVESLSG\SOT
1			)			TQEL\ENVGAHCPGCLRLLRAL\RMLKAWGR
						HSTGLRSLGMTITQCYEEVGLLLLFLSVGISIF
						STVEYFAEQSIPDTTFTSVPCAWWWATTSMT
			[ ]			
1						TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI
	· ·		i i			AIINDRFSACYFTLKLKEAAVRQREALKKLTK
			1			NIATOSYISVNLRDVYARSIMEMLRLKGRER
						ASTRSSGGDDFWF
828	2178	A	6786	5672	1360	GTHPASSGPVPLPPAAVSAATREELGEPVPFV
						TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ
	•		1			EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS
1					•	ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK
}						NGAGNGAKGKGKRERSISADSFDORDPGTPN
1						DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT
						APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS
i i			!	i	l	TEMANKAAEAVLKGQVETIVSFHIONISNNK
1 1						
					ļ	TERSTAPLNTQISALRNDPKPLPQQPPAPANQ
1		1		i	ļ	DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE
			i /		ļ	SPGVENKLIPSVGSPASSTPLPPDGTGPNSTPN
1 1				ŀ	ļ	NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT
, · I			1			LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP
1 1		- 1	1		j.	DEKEFTGAQSGGPQQNPGVLDGPQKKPEGPI
] 1			1	1		QAMMAQSQSLGKGPGPRTDVGAPFGPQGHR
1						DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP
	ļ		l			EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ
)		J				DMMVHQHGPRGVVRGPPPPYQMTPSEGWAP
	j	ļ	ĺ			
j [	1	,	l			GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ
1 1	ļ	j	ļ	j	1	MRLPGFAGMINSEMEGPNVPNPASRPGLSGV
1	ļ	[	1	ļ	1	SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP
				i	!	NPQGLSEEMFQQQLAEKQLGLPPGMAMEGIR
1	]	ļ.	İ		1	PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP
1	i	Ì				LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM
1 1	Į.	1	ľ	l	l	KGDVNLNVNMGSNSQMIPQKMREAGAGPEE
	i	1		1	Ī	MLKLRPGGSDMLPAQQKMVPLPFGEHPQOE
			ļ	1		YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDO
	1	- (	i	ľ	1	RTNSRLSHMPPLPLNPSSNPTSLNTAPPVORG
1		i		l	Í	
	- 1	- 1		0.	i	LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ
	i			1		SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA
		j	J	1	İ	SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS
.	i	4				
	ĺ	1		ĺ	1	PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG
						PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPPTASQPASVNIPG\SLPSSTPYTMPPEPTL
	*					

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Į.	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		t	9	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion, \=possible
	ļ		<u> </u>	sequence		nucleotide insertion
				· · · · · ·	}	KTVASSDDDSPPARSPNLPSMNNMPGMGINT
1	1	l	1	ł		QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ
1	İ	1				MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH
		i				GSQEPPMVPQGRMGFPQGFPPVQSPPQQVPFP
			1			HNGPSGGQGSFPGGMGFPGEGPLGRPSNLPQ
1		1		ŀ		SSADAALCKPGGPGGPDSFTVLGNSMPSVFT
1	j	l	1	ļ	•	DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ
		i .				YFPRGEVPGRKQPQGPGFGFHMQGMMGEQ
1		i				APRMGLALPGMGGPGPVGTPDIPLGTAPSMP
i						GHNPMRPPAFLQQGMMGPHHRMMSPAQST
	1					MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT
829	2179	A	6797	433	3	HPGPVGSPGMMMSMQGMMGP\NRTS ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH
023	21/3	^	1610	433	*	
	1					GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP
						ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG
		1		İ	-	AAASSTSSGASCPPVPASSRWGVRSRTRSGSG
830	2180	A	6800	3	1911	GEREPRDRPSERPRLV LPERAFGPRTPRAPRRRRRLLLSPPPRPPPPL
630	2100	Ι^	0800	· .	1911	DREPRAPGPWLCPSRAGTAODPARIRERRGR
1						VAGGAAGPAMELRARGWWLLCAAAALVAC
		ļ				
1		1	ŀ			ARGDPASKSRSCGEVRQIYGAKGFSSS\DVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA
1	ĺ	ĺ				ELETALRDSSRVLQAMLATQLRSFDDHFQHL
1						LNDSERTLQATFPGAFGELYTQNARAFRDLY
		i	1			SELRLYYRGANLHLEETLAEFWARLLERLFK
1	1	Į .		-		QLHPQLLLPDDYLDCLGKQAEALRPF\GEAP\
		l		*		RELRLRAT\RA\FVAAR\SFVQGLGVAS\DVVR
1	1	l	1			KVAQVPLG\PEC\SRAVIEAGSYC/ALHCVGVP
				i		GARPCPDYCRNVLKGCLANQADLDAEWRNL
İ	1					LDSMVLITDKFWGTSGVESVIGSVHTWLAEA
			ļ			INALQDNRDTLTAKVIQGCGNPKVNPQGPGP
						EEKRRRGKLAPRERPPSGTLEKLVSEAKAOL
1 .	1	ļ	1			RDVQDFWISLPGTLCSEKMALSTASDDRCWN
1						GMARGRYLPEVMGDGLANOINNPEVEVDIT
						KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF
			i i			QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS
			(8)			SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL
		1				LPLLLFLALTVARPRWR
831	2181	Α	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA
						EGRLREKLFSGYDSSVRPAREVGDRVRVSVG
-						LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS
						WDPAEHDGIDSLRITAESVWLPDVVLLNNND
]			j			GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS
						IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT
		ľ				GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS
			[ ]	ĺ		RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY
			1 1	J		LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF
						ALLTLTVFLLLLADKVPETSLSVPIIIKYLMFT
						MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV
				İ		RQIFIHKLPLYLRLKRPKPERDLMPEPPHCSSP
						GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL
]			j			SAPDLRRFIDGPNRAVALLPELREVVSSISYIA
					·	RQLQEQEDHDALKEDWQFVAMVVDRLFLW
						TFIIFTSVGTL\VIFLDATYHLPPPDPFP
832	2182	Α	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK
						ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY
						DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN
1						
į '			i i		. (	GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI
						GTDETLEVHDFKNGYTGIYFVGLQKCPIKTQI KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYWINPTL\IS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEEEGEDLHFP ANEKKGIEQNEQWVVPQVKVEKTRHARQAS EEELPINDYTENGIEFDPMLDERGYCCIYCRR
833	2183	Α	6846	116	602	GNRYCRRVCEPLLGYYPYPYCYQGGRVICRV IMPCNWWVARMLGRV EAEGEQVCGAKCCGDAPHVENREEETARIGP GVMESKEERALNNLIVENVNQENDEKDEKE QVANKGEPLALPLNVSEYCVPRGNRRRFRVR QPILQYRWDIMHRLGEPQARMREENMERIGE EVRQLMEKLREKQLSHSLRAVSTDPPHHDHH DEFCLMP
834	2184	A	6851	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSR GSREESPAPSRAPASASLWRRLVVVEAKMAA HAAAAAQAAAAQAAHAEAADSWYLALLGF AEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEA RTHLQLGSVLYHHTKNSEQARSHLEKAWLIS QQIPQFEDVKFEAASLLSELYCQENSVDAAKP LLRKAIQISQQTPYWHCRLLFQLAQLHTLEKD LVSACDLLGVGAEYARVVGSEYTRALFLLSK GMLLLMERKLQEVHPLLTLCGQIVENWQGN PIQKESLRVFFLVLQVTHYLDAGQVKSVKPC LKQLQQCIQTISTLHDDEILPSNPADLFHWLP KEHMCVLVYLVTVMHSMQAGYLEKAQKYT DKALMQLEKLKMLDCSPILSSFQVILLEHIIM CRLVTGHKATALQEISQVCQLCQQSPRLFSN HAAQLHTLLGLYCVSVNCMDNAEAQFTTAL RLTNHQELWAFIVTNLASVYIREGNRHQEVVL LYSLLERINPDHSFPVSSHCLRAAAFYVRGLF SFFQGRYNEAKRFLRETLKMSNAEDLNRLTA CSLVLLGHIFYVLGNHRESNNMVPPAMQLAS KIPDMSVQLWSSALLRDLNKACGNAMDAHE AAQMHQNFSQQLLQDHIEACSLPEHNLITWT DGPPPVQFQAQNGPNTSLASLL
835	2185	A	6855	315	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCL SVSGIGGFLVSLSSRMKLQTLAVSVTALKFWS AYVPCQTQDRDALRLTLEQIDLIRRMCASYSE LELVTSAKALNDTQKLACLIGVEGGHSLDNS LSILRTFYMLGVRYLILTHTCNTPWALSSAK GVHSFYNNISGLTDFGEKVVAEMNRLGMMV DLSHVSDAVARRALEVSQAPVIFSHSAARGV CNSARNVPDDILQLLEEERWAFVMVSLFHGE LIQWQPIRPMCSTVADHFDHIKAVIGSKFIGI GGDYDGAGKYRKKTTCKAPWRTSSRMSS PPRSRPSCWRKKVGPGRPWWWGGTGPPGQG RPEIRLLPLPMTGACGAVAASRTGSSGPG/SSL PNGHGGKGSGLANGLAGNP\GHLGLGSSFGT
837 .	2187	A	6863	2	1615	GPGSGRPPP  VLRGQRGPAGGLAEERRRGRNEWRIHDVTT APFPGLVQRRSRLLIVSQVRYFLKNKVSPDLC NEDGLTALHQCCIDNFEEIVKLLLSHGANVN AKDNELWTPLHAAATCGHIDLVKILVQYGA DLLAVNSDGNMPYDLCEDEPTLDVIETCMAY QGITQEKINEMRVAPEQQMIADIHCMIAAGQ DLDWIDAQGATLLHIAGANGYLRAAELLLDH GVRVDVKDWDGWEPLHAAAFWGQMQMAE LLVSHGANLNARTSMDEMPIDLCEEEEFKVL LLELK'HKHDVIMKSQLRHKSSLSRRTSHRQA S/SVGKVVRRTQPVGTGPNL\YRKEYE/GEEAI LWQRSA\AEDQRTSTYNGDIRET\RTDQENKD

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PNPRLEKPVLLSEFPTKIPRGELDMPVENGLR APVSAYQYALANGDVWKVHEVPDYSMAYG NPGVADATPPWSSYKEQSPQTLLELKRQRAA AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI EEMEEKVHGCCRIS
838	2188	A	6865	6291	739	AGPLEPRVQGAMALQLWALTLLGLLGAGAS LRPRKLDFFRSEKELNHLAVDEASGVYYLGA VNALYQLDAKLQLEQQVATGPVLDNKKCTP PIEASQCHEAEMTDNVNQLLLVDPPRKRLVE CGQLLKGNCALRALSNISLRLFYEDGSGEKSF VASNDEGVATVGLVSSTGPGGDRVLFVGKG NGPHDNGIIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQQFVAAFEDGPYVFFVFNQD KHPARNRTLLARMCREDPNYYSYLEMDLQC RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGIRITLGSNLGVQ AGDIQRISVAGRNCSPQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTQQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHIDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVYDTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVYDTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVYDTFFLPFSKDGDKDVMITGKL DIPPPRRPVVEQALYQFSNLLNSKSFLINFIHT LENGPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQODLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE RTKAITEIYLTRLLSVKGTLQQFVDNFFQSVL APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL YAKEISTYKKMVEDYYKGIRQMVQVSDQDM NTHLAEISRAHTDSLNTLVALHQLYQYTQKY YDEINALEEDPAAQKMQLAFRLQQIAAALE NKVTDL
839	2189	A	6872	1	1485	RARRLALQCHVCVCALTPGEQSGRRLPGQT WLMFSCFCFSLQDNSFSSTTVTECDEDPVSLH EDQTDCSSLRDENNKENYPDAGALVEEHAPP SWEPQQQNVEATVLVDSVLRPSMGNFKSRKP KSIFKAESGRSHGESQETEHVVSSQSECQVRA GTPAHESPQNNAFKCQET\VRL\QPRIDQRTAT SPKDAFETR\QDLNEEEAAQVHGVKDPAPAS TQSVLA\DGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRLLLYHITDGDNPTAVRHGCSL/F SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPEVLKWMNDLAKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLLCTDV
		A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKAINRN YLKQPVVKEKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRADAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENGFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNIRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKILEVHPERESNQEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSNDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHAFGDSKVTVNSIKPSSPTEGKRKIII
841	2191	A	6874	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methiorine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						KGLVQVTKHRLCRLPPSRAHLPTKEASSLHA VRTAPTSKVIKTRYRIVKKTPASPLSAPPFPLS LPSWRARRLSLSRSLVLNRLRPVASGGGKAQ PGSPWWRSKGYRCIGGVLYKVSANKLSKTSG QPSDAGSRPLLRTGRLDPAGSCSRSLASRAVQ RSLAIIRQARQRREKRKEYCMYYNRFGRCNR GERCPYIHDPEKVAVCTRFVRGTCKKTDGTC PFSHHVSKEKMPVCSYFLKGICSNSNCPYSHV YVSRKAEVCSDFLKGYCPLGAKCKKKHTLLC PDPARRGACPRGAQCQLLHRTQKRHSRRAAT SPAPGPSDATARSRVSASHGPRKPSASQRPTR 'QTPSSAALTAAAVAAPPHCPGGSASPSSSKAS SSSSSSSSPPASLDHEVAPSLQEAALAAACSNR LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDS GKPLHIKPRL
842	2192	A	6898	506	2071	WPDLVHTWSSEEAMGSCCSCPDKDTVPDNH RNKFKVINVDDDGNELGSGIMELTDTELILYT RKRDSVKWHYLCLRRYGYDSNLFSFESGRRC QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTFTTPGFAA QNLPNGYPRYPSFGDASSHPSSRHPSVGSARL PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDAP SVNKLVYENINGLSIPSASGVRRGRLTSTSTSD TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIR RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTPT TPLPQTPTRRTELYAVIDIERTAAMSNLQKAL PRDDGTSR/KTRHNST/DLPL
843	2193	A	6919	2	663	AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKHF QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM LEEGSFRGRTADFVFMFLFGGFLMTLFGLFVS L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRP SQLPGPLSALGAHGIFLVVGELNHCGPFGYCS WTHIFFLGRCISQSTWWNKNSENTIYFESYF
844	2194	A	6928	902	366	HRLCMPIQGACGERME/FSLLLPGLECNGVIL AHCNLRLPGSSNSPASASQVAGITGVCHHAR LIFVFSVETGFLHAGQAGLELLTSGDPPASAS QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845	2195		6939	1660	317	LYPENLGESLFPILLLPPPWPDGGRPCCVEMS TRAKKLRIWRILEEKESVAGAVQTLLLRSQE GGV\TSAAASTLSEPPRRT\QESRTRTRALGLPT LPMEKLAASTEPQGPRPVLGRESVQVPDDQD FRSFRSECEAEVGWNLTYSRAGVSVWVQAV EMDRTLHKIKCRMECCDVPAETLYDVLHDIE YRKKWDSNVIETFDIARLTVNADVGYYSWR CPKPLKNRDVITLRSWLPMGADYIIMNYSVK HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVIT YLAQVDPKGSLPKWVVNKSSQFLAPKAMKK MYKACLKYPEWKQKHL\PHFKPWL\HPEQSP LPSLALS\ELSVQHADS\LENIDESAV\AESREE R\MGGAGGEG\SDDDTSLYAEAPHRFRETETG PGAGRALGAAAAPALSPLHPPGTWWHRARP RRVLQPGWTEPQ

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghttamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
846	2196	A	6944	sequence 42	2672	mucleotide insertion  RRKMAGCRGSLCCCCRWCCCCGERETRTPE ELTILGETQEEEDEILPRKDYESLDYDRCINDP YLEVLETMDNKKGRRYEAVKWMVVFAIGV CTGLVGLFVDFFVRLFTQLKFGVVOTSVEECS QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE AGSGITEGKCYLYARQVPGLVRLPTLLWKAL GVLLTVAAMLLIGLGSPMHSGSVVGAGLPQ FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV VMGVIGGLLGATFNCLNKRLAKYRMRNVHP KPKLVRVLESLLVSLVTTVVVFVASMVLGEC RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV TLALFFVLYFLLACWTYGISVPSGLFVPSLLC GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA
847	2197	A	6951	3	1994	AFLGGVVRMTISLTVILIESTNEITYGLPIMVT LMVGKWTGDFFNKGNYDIHVGLRGVPLLEW ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL RNMCDEHIASEEPAEKEDLLQQMLERRYTPY PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ LVTLLVRGVCYSESQSSASQPRLSYAEMAED YPRYPDIHDLDLTILNPRMIVDVTPYMNPSPF TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE IVGIITRHNLTYEFLQARLRQHYQTI NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER
			·			LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST VGKRKIDQEGRVFQEKWERAYFFVEVQNIST CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY MERMRDEKLHELKKGLRKYLLGLSDTECPE QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK NFCINWSKLVSVASTGTPPMVDANNGLVTKL KSRVATFCKGAELKSICCIIHPESLCAQKLKM DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL WETHLTRNNLAHFPTLKLVSRNESDGLNYIP KIAELKTEFQKRLSDFKLYESELTLFSSPPSTKI DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
848	2198	A	6985	3	289	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C CFNAINTKIPIORLESYTRITNIOCPKEAVM
849	2199	A	6999	963	5	LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHLVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYVAICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA VLKLACADTHINENMVLAGAISGLVGPLSTIV VSYMCILCAILQIQSREVQRKAFCTCFSHLCVI

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence		/=possible nucleotide deletion, \=possible nucleotide insertion
						GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS
850	2200	Α	7001	1	1011	LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
	2500	••		· .		DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG
						KVARRRVGATWLLHLAVADLLCCLSLPILAV
						PIARGGHWPYGAVGCRALPSIILLTMYASVLL LAALSADLCFLALGPAW\CLRFS/GACGVQVA
						CGAAWTLALLLTVPSAIYRRLHQEHFPARLQ
						CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
1					•	YHLLGLVLTVAAPNSALLARALRAEPLIVGL
						ALAHSCLNPMLFLYFGRAQLRRSLPAACHW
061	2201		7011		0210	ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011	1	2310	AAASPLRMSRKGPRÄEVCADCSAPDPGWÄSI SRGVLVCDECCSVHRSLGRHISIVKHLRHSA
						WPPTLLQMVHTLASNGANSIWEHSLLDPAQV
						QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF
						VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG
						QTLQAELLVVYGADPGSPDVNGRTPIDYARQ
						AGHHELAERLVECQYELTDRLAFYLCGRKPD
<b>!</b>	:					HKNGHYIIPQMADSLDLSELAKAAKKKLQAL
						SNRLFEELAMDVYDEVDRRENDAVWLATQN HSTLVTERSAVPFLPVNPEYSATRNQGRQKL
						ARFNAREFATLIIDILSEAKRRQQGKSLSSPTD
						NLELSLRSQSDLDDQHDYDSVASDEDTDQEP
						LRSTGATRSNRARSMDSSDLSDGAVTLQEYL ELKKALATSEAKVQQLMKVNSSLSDELRRLQ
						REIHKLQAENLQLRQPPGPVPTPPLPSERAEH
1.						TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG
						PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK
						LSRHGSGADSDYENTQSGDPLLGLEGKRFLE
•						LGKEEDFHPELESLDGDLDPGLPSTEDVILKT
						EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA
						VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
<u> </u>						KAAKQLVTITTREKKQ
852	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL
						TVKGLLKPSFSPRNYKALSEVQGWKQRMAA KELARQNMDLGFKLLKKLAFYNPGRNIFLSP
] ]			*			LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP
						EKDLHEGFHYUHELTQKTQDLKLSIGNTLFID
				*		QRLQPQRKFLEDAKNFYSAETILTNFQNLEM AOKOINDFI/ESKTHGKINNLIENIDPGTVMLL
.						ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS
] .						VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK
						NITAIFILPDEGKLKHLEKGLQVDTFSRWKTL LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS
j )						KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM
						DERGTEGAAGTGAQTLPMETPLVVKIDKPYL
853	2203	A	7017	1 .	3293	LLIYSEKIPSVLFLGKIVNPIGK MTHACNPSTLGGQGRRITRSHGRRRSSRGPV
6.5	. 2203	^	/01/		3473	ARHVAAGAGHENKHGGSRRFPAGVAPRRAM
						ANVSKKVSWSGRDRDDEEAAPLLRRTARPG
						GGTPLLNGAGPGAARQSPRSALFRVGHMSSV
0.0			j			ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY ESLDYDNSENQLFLEEERRINHTAFRTVEIKR
]						WVICALIGILTGLVACFIDIVVENLAGLKYRVI
			·			KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG ASFWNQFLTWRIFFASMISTFLNFVLSIYHG NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI AMGVVGGVLGAVPNALNYWLTMFRIRYIHIR PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL QGSM\$YPLQLFCADGEYNSMAAAFFNTPEK SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT VIMMEATSNVTYGFPIMLVLMTAKIVGDVFIE GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV MSTPVTCLRRREKVGVIVDVLSDTASNHNGF PVVEHADDTQPARLQGLILRSQLIVLLKHKVF VERSNLGLVQRRLRLKDFRDAYPRFPPIQSIH VSQDERECTMDLSEFMNPSPYTVPQEASLPR VFKLFRALGLRHLVVVDNRNQVVGLVTRKD LARYRLGKRGLEELSLAQTGPKAQATAEGRV AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP LSLEELSERYESSHPTSTASVPEQDTAKHWNQ LEQWVVELQAEVACLREHKQRCERATRSLL RELLQVRARVQLQGSELRQLQQEARPAAQAP EKEAPEFSGLQNQMQALDKRLVEVREALTRL
				į		RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ QEEQGREVACGALQKNQEDSSRRVDLEVAR M
854	2204	A	7037	139	2604	AGTWEPRPYDQAKETGAPGSQPPVPPMELRP WLLWVAATGTLVLLAADAQGQKVFTNTW AVRIPGGPAVANSVARKHGFLNLGQIFGDYY HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL EQQVAKRTKRDVYQEPTDPKFPQQWYL\SG VTQRDLMVKAAWAQGYTGHGIVVSILDDGI EKNHPDLAGNYDPGASFDVNDQDPDPQPRY TQMNDNRHGTRCAGEVAAVANNGVCGVGV AYNARIGGVRMLDGEVTDAVEARSLGLNPN HIHIYSASWGPEDDGKTVDGPARLAEEAFFR GVSQGRGGLGSIFVWASGNGGREHDSCNCD GYTNSIYTLSISSATQFGNVPWYSEACSSTLA TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS APLAAGIIALTLEANKNLTWRDMQHLVVQTS KPAHLNANDWATNGVGRKVSHSYGYGLLD AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAQARLT LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEOTCSROS
		į				QSSRESPPQQOPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	1441	QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF LVKEKGYDKELLNVTPEDWDFCCKGLALDL EDGNFLKLANNGTVLRASHGTKMMTPEVLA EAYGKKEWKHFLSDTGMACRSGKYYFYDN

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEALPSLDKPGWYSQ GNAVHLYELLKKMTGKPEPKVVYFGDSMHS DIFPARHYSNWETVLILEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFFU DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV LSSDETLISK
856	2206	A	7082	396	1635	SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRFCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQLQSLESKLT SVFFMGDMGSFEEDRINATVWKLQPTAGLQ DLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYTTNLGLFSLVFLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLG\LP WALIFFSFASGTFQLVVLYLFSITISFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3 ,	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO:	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	- Cuitoc	•	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
10000		1	7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		1	1	peptide		/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
		l —			f	HPPPEEDQGEERPRLWVMPNHQVLLGPEEDQ
			1		1	DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP
1		l				CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI
-1		1				LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL
ì						KSAMILQ .
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF
						YKNIQKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP
1		l				LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA
1		l	<b>{</b>		ľ	RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ
1		1.				CIKPNYQSPPKECDYNILANSVA
864	2214	A	7214	845	1619	SDKGGKKADRKNHLRHAFPLLPHRVRERLH
						DPKVPVDADHVQGQDPGRAAHDIHGEDVTE
		1				KVSKDPLAPDEVGDTDEGHDRHGHREVGQR
J		l	]			HGHDQEEVAYEERACEGGKFATVEVTDKPV
		ł				DEALREAMPKVAKYAGGTNDKGIGMGMTV
		l				PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP
						PAPSDKSVKIEEREGITVYSMQFGGYAKEAD
						YVAQATRLRAALEGTATYRGDIYFCTGYDPP
ļ		Ĺ <u>.</u>				MKPYGRRNEIWLLKT
865	2215	Α	7246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS
0.55	0016				4010	LANMAKPRLY
866	2216	Α	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL
1			1			DLKKSDFSTRWQKQRCPVVKSKCRENASPFF
1						FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV
1 .		1	i i			QIPLTESYCGPCPKNWICYKNNCYQFFDESKN
						WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT
		l	]			IIEMQKGDCALYASSFKGYIENCSTPNTYICM
1	•					ORTV
867	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFROOVVFI
00,	2211	^	1200	151	370	MPFFQTLWLMNANRFCSIFTTTNVANNCWW
			1 1			TPYHCWLSVVVCRCESHGI
868	2218	Α	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP
'				-		KGSCPAGGSRMVSESD*EGRGC*ASYPCAC*
						AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	A	7332	1223	332 .	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR
						FLTLCTWLLLLGPGLLATVRAECSQDCATCS
						YRLVRPADINFLACVMECEGKLPSLKIWETC
						KELLQLSKPELPQDGTSTLRENSKPEESHLLA
						KRYGGFMKRYGGFMKKMDELYPMEPEEEA
						NGSEILAKRYGGFMKKDAEEDDSLANSSDLL
						KELLETGDNRERSHHQDGSDNEEEVSKRYGG
			-			FMRGLKRSPQLKEKAKELQKRYGGFMRRVG
1						PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE
						SYSKEVPEMEKRYGGFMRF
870	2220	Α	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG
						AGQGRGREGAESGGSRGEGPGSDGRLPATGD
						FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT
						TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL
1 1						KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF
						VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD
1 1						FYITLGTGCVFLLASIIFVSTHDRTSAEIAAIVF
						GFIASFMFLLDFITMLYEKRQESQLRKPENTT
071	2001		7402		202	RAEALTEPLNA
871	2221 .	Α	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN
					1	DPGNMSFVKETVDKLLTGFRCFREREAAPRR
L		نــــــا	L l			ALRGAALPGESEAGDPESLRSSVNADWIQYS

No. of   N	SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
			nou .				
1914   ng to first amino aaid residue of peptide peptide sequence							
mino aed residue of peptide requence peptide sequence pep	seq-	uence	1	09/496		to last amino	
Peptide sequence	uence			914			
Peptide   Sequence							
	İ					sequence	
DLWEAEVSTRECAGFCQSCFRTFGNQEKDG   PRIC			1				
FFIC	_			-			
PGGS*PQATI-HI_DRARVSASPTKEIQVAKVV   CGLIKPCPANYTAKICSGAANVVGPTMCPED			1				
CGLIKPCPANYFAKICSGAANVVGPTMCEN	872	2222	A	7413	1061	359	
RMIMSPYKNNYGRQLNIALVAGTGALVLQA   KAPDMYSGDVMILVFKLEREGGALVLVAQ   VDDPGTKMDESRKLFSDLGSSYAKQLGFGA   SWYPIGAKDLRGKSPFEOFLKEQPQTQNKYE   GWPELLEMEGCMPPKDF     B73   2223   A   7429   2242   2394   ILKCAGHIGGSCL*SQHFGRLRWEDRLRLGVQ   DHEGGHCSTIFSLKERKLF							
STOPETIKMNDESRIKLFSDLGSSYAKQLGFRD	i						
SWYFIGAKDLRGKSPFEQFLKEQPQTÖNKYE							
873   2223   A   7429   2242   2394		}	<u> </u>	ĺ			
DHPGGHCETTSLLKIERKLF							
874   2224   A   7468   146   894   PCTSCVLWATLHLPASTRKAPQAECGMISTIE	873	2223	Α	7429	2242	2394	,
WQKIGVGITGFGIFFILFGTILLYFDSVILAFGN   LIFLTGLSLIGLRKTFWFFFQRHKLKGTSFLL   GGVVTVLLRWPLLGMLEHYGFFSLFKGFFPV   AFGFLGNVCNIPFLGALFRRLQGTSSMV-KTE   MSSLNLDHWLKGAKREEWEPPPQSPALTHSP   TYFGPPQVQKERNGABQLTSNPQVDSRGCQB   AEMQTPRRLQGWGWYHTLY-WEEK   AEMQTPRLGWGWYHTLY-WEEK   AEMQTPRLGWGWYHTLY-WEEK   AEMQTPRLGWGWYHTLY-WEEK   AEMQTPRLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WEEK   AFGFLGWGWYHTLY-WARPW   LGQGGWA*QGFKTSLGNTVRQCYY   LGCAPDTRFYPEPGGGAAPWVALVARGC   TFLDKVLVAARRNASAVVLNTERY-GNTT   MSHAGTGKNIVMISYFYKGREILELVQKOIPV   TMTIGVGTRHVQEFISGGSVVFVALAFITMMI   SLAWLIFYYIQRFLYTGSQIGSQNFKETKKVI   GQLLHTVKHGEKGIDVDAENCAVCIENFKV   KDIIRLPCKHIFHRICIDPWLLDHRTCPMCKL   DVIKALGYWGEPGDVQEMPAPESPPGRDPA   NLSLALPDDDGSDSSPSSPSASPASEPCOPSF   KGDAGENTALLEAGRSDSHGGPIS	974	2224	_	7460	146	904	
LLE.T.GLS.I.IGLRKITFWFFFQRKILKGTSFLL   GGVVIVLIRWPLLGMELETYGFFSLKEGFPV   AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE   MSSLNLDHWLKGAKREEWEPPPGSPALTHSP     TYPGPPQVQKERNGAGQLTSNPQVDSSGCQE   AEMQTPRRLGWGWYHTLTLYLWEEK     875   2225   A 7498   91   251   GEKPVPTVLQDEAGQULTSNPQVDSSGCQE   AEMQTPRRLGWGWYHTLTLYLWEEK     876   2226   A 7544   403   587   YSCLCFLFKHITISFKNSYHIWLGTVAPQWGWPG     877   2227   A 7566   2   940   GCAPDTRFFVPEPGGRGAAPWVALVARGGC   TFKDKVLVAARRNASVLVINEERY GONTEN     878   2227   A 7566   2   940   GCAPDTRFFVPEPGGRGAAPWVALVARGGC   TFKDKVLVAARRNASVLVINEERY GONTEN	6/4	2224	_ A	/408	140	094	
GGVVIVLIRWPILGMFLETYGFFELKGFFPY AFGLGNYCNIPFLGALFRILQGTSSMV*KTE MSSLNLDHWLKGAKREWEPPPQSPALTHSP TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE AEMQTIPRILGWGYVTLTLTLVLWEEK 875							
MSSLNLDHWLKGAKRREEWEPPPQSPALTHISP						1	•
TYPGPPQVQKERNGAEQLTSNPQVDGRGCCE			1				AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
AEMQTPRRGWGWYHTLTLYLWEEK	- 0			!			
875	ľ		1				
SERHEP*HGÖVLFRLGPSAPPGKL	875	2225		7498	91	251	
876	0.5		l "	7-170	٠.	~3.	
877   2227   A   7566   2   940   GCAPDTRFFVPEPGGRGAPWVALVARGGC	876	2226	Α	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
TFKDKVLVAARRNASAVVLYNEERYGNITLP MSHAGTGNIVVIMISYPKGREILELVQKGIPV TMTGVGTRHVQEFISGGSVVFVAIAFITMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDSSPFSASPASEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSKSTLKHIWTESSKDLSISRILS QTFRGKENDTDLDLRYDTFEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTYFFSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG B80 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGFG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDLALGPILKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITTSVSSWPYSSHRMRFIT NHSDQPPONFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIALLYVFLGHRKRNSIQIYL LNVAIADLLLIFCLFFRMYHDNQNKWTLGVIL CKVVGTLFRILVENNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
MSHAGTGNIVVIMISYPKGREILELVQKGIPV TMTIGVGTRHVQEFISQSVVFVAIAFITMMII SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIIRLPCKHIFHRICIDFWLLDHRTCPMCKL DVIKALGYWGEPGDVOGEMPAFESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSILCKVDVRWITYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSKSSTLKHHWTESSKDLSISRILS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLINLLITGKIVDHGNGTFSVYF RHNSTGQGNSVSVLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQDSLYPVGYLDKQVPDTS VQCTDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLINGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTISVSSWPYSSHRMFFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIALLYVFLGHRKRNSIQIYL LNVAIADLLLIFCLFFRMYHDNQNKWTLGVIL CKVVGTLFTRINGNINGLILGFISLDRYIKINRSI QQRKAITIYKQSIVVCCIVWMLALGGFLTMIIL	877	2227	A	7566	2	940	
TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII SLAWLIFYYIGRLYTGSQIGSQSHRKETIKKVI GQLLLHTVKHGEKGIDVDAENCAVCIENFKV KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAFESPPGRDPAA NLSLALPDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSILCKVDVRWIYVSEGTKTQRHRQGSLR RGRMQAACWYVIFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTI.KHIWTESSKDLSISRILS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQERPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLINLITIGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSTCLDQLNGSAPGNVFHG 880 2230 A 7612 93 659 DAAVAMTAQGLVANRGRFFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSKFLQGLVYLIGNLMGLALAVYK CQSMGLIPTHASDWLAFTEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFLIFTNAMMYISILLGFISLDRYIKINRSI QQRKAITTKQSIVVCCIVWMLALGGFLTMIIL							l '
SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI GQLLLHTVKHGEKGIDVDABNCAVCIEDIFKV KDIIRILFOKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPFSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRFHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFORCHEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG B80 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRFFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							· · · · · · · · · · · · · · · · · · ·
KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL DVIKALGYWGEPGDVQEMPAPESPPGDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTDQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICTISFYSTD YKLVQKVCPDYNYHSDTYPFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG GGSRGRSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNNYSIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	İ				•		1
DVIKALGYWGEPGDVQEMPAPESPPGRDPAA NLSLALPDDDGSDESSPPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNILALYVFLGHRRRNSIQIYL LNVALADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL			}			}	
NLSLALPDDDGSDESSPSASPAESEPQCDPSF KGDAGENTALLEAGRSDSRHGGPIS  878 2228 A 7586 315 1232 ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCREYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVALADLLLIFCLPFRIMYHNQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
KGDAGENTALLEAGRSDSRHGGPIS							
878   2228   A   7586   315   1232   ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR RGRMQAACWYVLFLLQPTVYLVTCANLTNG GKSELKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLRYDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG   879   2229   A   7605   479   391   TESWKLKWWSPTCLDQLNGSAPGNVFIHG   880   2230   A   7612   93   659   DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSQQDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLPTHASDWLAFIEPPERMEFSGGGL LL   1452   SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1		,				_
GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS QTFRGKENDTDLDLR YDTPEPYSEQDLWDW LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	878	2228	Α	7586	315	1232	
RESIDENT A TOTAL OF THE PROPERTY OF THE PROPER							
LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVTL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL			ľ				
GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVTL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1						
RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG  879 2229 A 7605 479 391 TESWKLKWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRFFKWAIELSGPG GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVTL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL					•		
AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD YKLVQKVCPDYNYHSDTPYFPSG 879 2229 A 7605 479 391 TESWKLKWSFTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRFKKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL 881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	-						
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	1						AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK
879 2229 A 7605 479 391 TESWKLKWWSPTCLDQLNGSAPGNVFIHG 880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL						<b>.</b>	
880 2230 A 7612 93 659 DAAVAMTAQGGLVANRGRRFKWAIELSGPG GGSRGSDRGSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	970	2220		7604	470	201	
GGSRGRSDRĞSGQGDSLYPVGYLDKQVPDTS VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
VQETDRILVEKRCWDIALGPLKQIPMNLFIMY MAGNTISIFPTMMVCMMAWRPIQALMAISAT FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTITMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVTL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	""		**	7012		(3)	
FKMLESSSQKFLQGLVYLIGNLMGLALAVYK CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL		•					VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
CQSMGLLPTHASDWLAFIEPPERMEFSGGGL LL  881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
881 2231 A 7615 291 1452 SPQKTMRSHTTIMTTTSVSSWPYSSHRMRFIT NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT SYSVIFTVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	881	2231	A	7615	291	1452	
SYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYL LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL		•				.	
CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL							
QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL	ŀ					· i	

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  VVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPN SGKYATTARNSFIVLIFTICYPYHAFRIYISS QLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDP VMYFLMSSNIRKIMCQLLFRRFQGEPSRSEST SEFKPGYSLHDTSVAVKIOSSSKST
882	2232	A	7617	67	379	RQMALLKANKOLISAGLKEFSVLLNQQVFND PLVSEEDMVTVVEDWMNFYINYYRQQVTGE PQERDKALQELRQELNTLANPFLAKYRDFLK SHELPSHPPPSS
883	2233	A	7622	400	215	KVKTCRYNPKYSAANDTGFVDIPSREKDLAK
884	2234	A	7638	2640	2861	AVATVGPISVAVGASHVFFQFYKKGKHLSS APVLILQMVKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRKVSECRQMGPGALEQFP GLSCHTSHSG
885	2235	A	7642	201	455	PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFFTAWFFVYEVTSTKYTRDIYKELLISL VASLFMGFGVLFLLLWVGIYV
886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK
887	2237	A	7693	85	315	NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK
888	2238	A .	7702	242	1298	APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC
889	2239	A	7707	185	2911	CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGEL WSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

2241   A   7721   61	SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide scq- ucnoc	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phemylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP DESTKTKDQILTSRINAVERDLLEPSPADQLG NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV TREPARRLFLFGEEPSKLDQDVLAALECADV DPHQFPAVHRWKSAVLCYSFSDRQSWPSPAV KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP GLGSPGRYSPVHGSQLRRMARLAELAAL
WSQYEKLDAGEORLMNEAFCPASDLFGFTTI	890	2240	A	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEELRAAVLQLRETVVQQ KETLASARARIELTGKLARCEGLAGGKARGA GATGKDTIMGDLPRDPGHVVEQLSRSLQTLK DRLESLEPLPAMPMQGGAQSPEELRAAVLQ LRETVYQQKETLASARARIELTGKLARCEGL AGGKARGAGTGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVVQQKLGELERQLLRKGGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT ICLWLRSSASPGMGTPFSYAVPQQANEIVLE WGNNPIELLNDKVAQLPLFVSDGK WHHCV TWTTRDGMWEAFQDGKKLGTGENLAPWHPI KPGGVLLIGGEQDTVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD NNVDVFGGASKWPVETCEERLLDL NNVDVFGGASKWPVETCEERLLDL DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVLLERHRQTRSSTET FLFHLAVADLLLVFILPFAVAGGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACLAVDRYLAIV HAVHAYRHRILSIHITCGTIWLVGFLLALPEI LFAKVSQGHLNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYYGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV HAVHAYRHRLSHNTCGTIWLVGFLLALPEI LFAKVSQGHLNNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYYGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV GLAHCCLNPMLYTFAGVKFRSDLSRLITKLG CTGPASLCQLFFSWRRSSLSSESENATSLTTF FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILRFVPVKMSFGLIRVFSI VIPPLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR LWLSLFLHAGKEAPHCPRTRPL					,		VSQYEKLDAGEQRLMNEAFQPASDLFGPITL HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS WNFVFGQASLTDGVGIFSFARYGSDFYSMHY KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL EEADRRPLNLCPICLHKLQCAVGFSIVERYKA LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV HAVHAYRHRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHINNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG CTGPASLCQLFPSWRRSSLSESENATSLTTF  894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR LWLSLFLHAGKEAPHCPRTRPL			A		2	1650	SAPTAPARPCRAERGSGGGMLALLAASVALA VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEEELRAAVLQLRETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPRDPGHVVEQLSRSLQTLK DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ LRETVVQQKETLASARAIRELTGKLARCEGL AGGKARGAGATGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVLQQRLGELERQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT ICL WLRSSASPGMGTPFSYAVPGQANEIVLIE WGNNPIELLINDKVAQLPLFVSDGKWHHICV IWTTRDGMWEAFQDGKKLGTGENLAPWHPI KPGGVLILGQEQDTVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D  895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR LWLSLFLHAGKEAPHCPRTRPL	893	2243	A	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHILAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAHCCLNPMLYTFAGVKFRSDLSRLLTKLG
LWLSLFLHAGKEAPHCPRTRPL							FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D
	895	2245	A	7753	119	278	
	896	2246	A	7754	1	372	LWLSLFLHAGKEAPHCPRTRPL SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	ucnœ	1	09/496 914	ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, O=Glutamine, R=Arginine, S=Serine,
ucuce	ł		714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	İ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	l	ł		peptide	sequence	/=possible nucleotide deletion, \=possible
1	İ	ł		sequence		nucleotide insertion
-	<del></del>	<del> </del> -	<del> </del>	Sequence		QVSTAALAVLLCTMALCNQVLSAPLAADTPT
	ł		İ			ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
ł						TKRGRQVCADPSEEWVQKYVSDLELSA
897	2247	A	7761	1725	445	RPRRGTHHFSCVLGSFRVSAMFPRVSTFLPL
1			'''	•••	'''	RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY
1	ľ	1				RSSALLARTKNNIQRYFGTNSVICSKKDKQSV
i		1	-			RTEETSKETSESQDSEKENTKKDLLGIIKGMK
ļ		ļ				VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA
	1	1	1	1	1	TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ
			1		l	TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI
	l	ł	ļ.		l	ISDMKVARSATARVRSRPELRIQFDEGYDNYP
ł						GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE
		l	ł		ļ	APETDTSPSLWDVEFAKQLATVNEOPLONGF
1	l	l			}	EELIQWTKEGKLWEFPINNEAGFDDDGSEFH
	Ì	1				EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP
			}			YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF
						KLRPWKFLFRNN
898	2248	Α	7775 '	85	496	SCQTTQPPAQSCSTGTMRIMLLFTAILAFSLA
						QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL
1	1					QRLFKSHSSLEGLLKALSQASTDPKESTSPEK
1	ł	ļ	}			RDMHDFFVGLMGKRSVQPDSPTDVNQENVP
						SFGILKYPPRAE
899	2249	Α	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI
[		1 '	1			FSKSMNESMKNQKEFMLMNARLQLERQLIM
						QSEMRERQMAMQIAWSREFLKYFGTFFGLA
]						AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL
		1				GYGTLLERMKGEAEDILETEKSKLQLPRGMIT
000		<u> </u>				FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE
			-			GRDKDTFSKMAMVSEFLKQAWFIENEEQEY
						VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH
1 1			1 1			KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY
1						LQETGKPLDETLKKALTGHLEEVVLALLKTP
1						AQFDADELRAAMKGLGTDEDTLIEILASRTN
					i	KEIRDINRVYREELKRDLAKDITSDTSGDFRN
				1	ļ	ALLSLAKGDRSEDFGVNEDLADSDARALYEA
				İ		GERRKGTDVNVFNTILTTRSYPQLRRVFQKY
j						TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA
1 1				ŀ		TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS
				ļ		RSEIDMNDIKAFYQKMYGISLCQAILDETKGD
901	2251	A	7796	2	807	YEKILVALCGGN VEFHPQRARAGARAPSMGVLLTQRTLLSLVL
		^	,,,,,	- 1	307	
						ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP
j						GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL
			[	j	1	ADLEQRLPKAQDLERSGLNIEDLEKLOMARP
					Ì	NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA
1	l		l ,	1	J	SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH
		į		.	l	SVGRVFSKWGESPNRSRRHSPHQALRKGVRR
						TRPSRKGKRLMTRGOLPR
902	2252	Ā	7802	2	721	TAARRQKGTAARRLQKGTAARRQKGTAA
		*	,502	-	'	RRRQKGTAARRPOKGTAARRROKGTAARRR
		. ]			İ	QKGTAARRQKGTAARRPQKGTAARRRQKG
				]		TAARROKGTAARROKGLAIASRGCPCASR
	j	I	J	ļ	J	AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD
	ļ	- 1		į		CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF
	j	ĺ	.		l	LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV
1	1	l	1	ł	]	REKPDDPLNYFLGGCAGGLTLGARTHNYGIG
	J	l	ļ	i	1	AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SEQ	ID   SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO:	of NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-			in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotid			USSN	location	corresponding	I=Isolaucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uenc	e	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ		]		peptide	j	/-possible nucleotide deletion, \-possible
000	2252		7007	sequence	504	nucleotide insertion
903	2253	A	7807	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMA
1						VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
ł	ì	ľ	1	1		RGDSPTDSQKDMIEIPLPPWQERTDESIETKR
1					•	ARLLYESRKRGMLENCILLSLFAKEHLQHMT
1						EKQLNLYDRLINEPSNDWDIYYWATEAKPAP EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
1	1	}	1	ł	ļ	
904	2254	A	7813	40	821	LFEKPR   GAGRALGHLETGAGDVAAALPARKFPRSLLG
304	2234	Α .	/613	40	021	AGARLTGWTMNVFRILGDLSHLLAMILLLGK
1		1				IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
1	1	1			ł .	ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF
İ	1	i	1	ì	ĺ	DSENDTFRLEFLLVPVIGLSFLENYSFTLLEIL
				Į		WTFSIYLESVAILPOLFMISKTGEAETITTHYL
			1			FFLGLYRALYLANWIRRYQTENFYDQIAVVS
		l	į	<b>!</b>	}	GVVQTIFYCDFFYLYVTKGRSWDDSNADTGL
1		l				RSYSSI
905	2255	A	7817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVOSFSEA
			''''			QTEMVRTLERKLEAKMIKEESDYHDLESVVQ
1			1 .		j	OVEONLELMTKRAVKAENHVVKLKOEISLL
1	į.	ł	ł			QAQVSNFQRENEALRCGQGASLTVVKQNAD
			ł			VALONLRVVMNSAQASIEOLVSGAETLNLVA
						EILKSIDRISEVKDEEEDS
906	2256	A	7822	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL
						LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG
	1.					HOPOTGSGESSGASGDKDHLYSTVCKPRSPK
1						PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA
ĺ						TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK
1			ł			RPSLPSSPSPGLPKASATSATLELDRLMASLSD
1						FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG
						KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
						NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
			1			GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
		,				RHKMVTALGTHWHPEHFCCVSCGEPFGDEG
1						FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
1	1 1			]		YISALSALWHPDCFVCRECFAPFSGGSFFEHE
				•		GRPLCENHFHARRGSLCATCGLPVTGRCVSA
1				j		LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY
						CQPCFLKLFG
907	2257	Α	7828	1792	1671	FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH
1		<u>_</u>				YCKSQAWG
908	2258	A	7842	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG
						SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET
	i i					YTCPKMIEMEQAEAQLAELDLLASMFPGENE
						LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI
						NMNLDVSDEKMAMFSLACILPFKYPAVLPEI
				. !	•	TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV
1			,			CILNATEWVREHASGYVSRDTSSSPTTGSTVQ
						SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL
						SLSGFSMPGKPGVVCVEGPQSACEEFWARLR
ŀ						KLNWKRILIRHREDIPFDGTNDETERQRKFSIF
			( 1			EEKVFSVNGARGNHMDFGQLYQFLNTKGCG
000	10000		7070	2000	-0000	DVFQMFLWV
909	2259	A	7870	3067	2923	EGICVYTFIYVHMYTRTCMHTYPYMYMNSV
010			500:		4054	LISSEILLIPSKYLFESK
910		A	7884	212	4874	GALTWSHPLLAVCPQGVWLGSTPSGSPALLP
(	2260	**	1	,		DOITO :
[	2260					PSHRVNAEPGCVVTNACASGPCPPHANCRDL
	2260				,	WQTFSCTCQPGYYGPGCVDACLLNPCQNQG
	2260				,	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLELNKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRPSPSEVLPTSSSIENSTTSS VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERRGARLPQNPVMNSPVVSVAVFHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGHANVAAA LGVAELLFLLGHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAA WMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA
						QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPFRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRRTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVFGLYSSFYPVFIY FLFGTSRHISVESLCVPGPVDT
911	2261	A	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL
	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL EILVKEDRDSGVNFQPEDTCARLRCSLHASLL VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH RHQLFINFRDLGWHKWIIAPKGFMANYCHGE CPFSLTISLNSSNYAFMQALMHAVDPEIPQAV CIPTKLSPISMLYODNNDNVILRHYEDMVVD
913	2263	Α	7892	15	849	ECGCG  ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS VGATCCYLLSSIFGKQLVVSFPDKVALLQR KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN RLSNITRPFFSKKKKKILV
915	2265	A	7909	3	641	HASGPGGLLRRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV WRVKFTLAKIRKAARELLTLDEKDPRRLFEG NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS HVKRKNASKGQGGAGARDDEEEE
916	2266	A	7914	3	967	VAHTOWHTCQRLSQLTHRSILKYLLIDTHAC QVLILKHTHASLSLPSCQECFPSSIPSASHMVS HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS SILPTGDSWGMLACLCTVLWHLPAVPALNRT GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA TAELRRSLAHFCTSLQGLLGSIAGVMAALGY PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL LKELQTWLWRSAKDFNRLKKKMQPPAAAVT LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED THHQIATLTVLVAPENPVVEVREQAVEGGEV ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ ENGKVWSVASTVRFRVDRKDDGGIIICEAQN QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL PERAEAVGETLTLPGLVSADNGTYTCEASNK HGHARALYVLVVYGESRLRPTEGGGGAPDP GAVVEAQTSVPYAIVGGILALLVFLIICVLVG MVWCSVRQKGSYLTHEASGLDEQGEAREAF LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPSSSVSSSLSPSAVVMACRWSTK ESPRWRSALLLFLAGVYGNGALAEHSENVH ISGVSTACGETPEQIRAPSGITSPGWPSEYPAK INCSWFIRANPGEIITISFQDFDIQGSRRCNLD WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR CGNGKCIPEAWKCNNMDECGDRSDEEICAKE ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VILRFTDFKLDGTGYGDYVKIYDGLEENPHK LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV NAARGFNATYQVDGFCLPWEIPCGGNWGCY TEQQRCDGYWHCPNGRDETNCTMCQKEEFP CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF CQPGNFHCKNNRCVFESWVCDSQDDCGDGS DEENCPVIVPTRVITAAVIGSLICGLLLVIALG CTCKLYSLRMFERRSFETQLSRVEAELLRREA PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA RSRHSGSLALVSADGDEVVPSQSTSREPERNH THRSLFSVESDDTDTENERRDMAGASGGVAA PLPQKVPPTTAVEATVGACASSSTQSTRGGH ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL ASDQGQGLRQPYNATNPGVRPSNRDGPCERC GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
920	2270	A	7953	47	572	GGWNDVACHTTMYFMCEFDKKNM GGRASWPEQAKEPRREGHTDKQQTEDVLAA GLRCLPHLPAICARRMSPAFRAMDVEPRAKG VLLEPFVHQVGGHSCVLLFFNETTLCKPLVPRE HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS FSSSWPSAQHLTPSVFNPW
921	2271	A	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF MMVYSFRALSFKESTWATFQHGGEATKSRSL SSTQ
922	2272	A	7967	1443	1660	ENITEK WKEIWMCRGNKKSCCWTFIKDRHLT VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN WELVKPN
923		A	7981	1	3023	GSAPRAATAMARARPPPPPSPPGLLPLLPLL LLPLLLLPAGCRALEETLMDTKWVTSELAWT SHPESGWEEVSGYDEAMNPIRTYQVCNVRES SQNNWLRTGFIWRRDVQRVYVELKFTVRDC NSIPNIPGSCKETFNLFYYEADSDVASASSPFW MENPYVKVDTIAPDESFSRLDAGRVNTKVRS FGPLSKAGFYLAFQDQGACMSLISVRAFYKK CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN AVEVSVPLKLYCNGDGEWMVPVGACTCATG HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP NSRTTSPAASICTCHNNFYRADSDSADSACTT VPSPPRGVISNVNETSLILEWSEPRDLGVRDD LLYNVICKKCHGAGGASACSRCDDNVEFVPR QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG IASTVTSQMNSVQLDGLRPDARYVVQVRART VAGYGQYSRPAEFETTSERGSGAQQLQEQLP LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL KQPGRREVFVAIKTLKVGYTERQRRDFLSEA SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM KYLSEMNYVHRDLAARNILVNSNLVCKVSDF GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI AYRKFTSASDVWSYGIVMWEVMSYGERPY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Animo aciu sequence (A=Alanine C=Cysteine,
		поа	1			D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ì	ł	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł		i	į .	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	peptide	Sofounce	/=possible nucleotide deletion. \=possible
	•		i			nucleotide insertion
	<b></b>	<u> </u>	<b></b> -	sequence		nucleonae insernon
1	ì	l	ł	1	i d	WDMSNQDVINAVEQDYRLPPPMDCPTALHQ
		1	l			LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA
1	l	i	1	ł		SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD
i		1	i	ĺ		WLDAIKMGRYKESFVSAGFASFDLVAQMTA
1	1			l	1	EDLLRIGVTLAGHQKKILSSIQDMRLQMNQT
ł	ł	l	l	ł	l	LPVQV
924	2274	A	7985	1	-503	
324	2214	^	7985	*	-505	FRPRTKKATAMYLEHYLDSIENLPCELQRNF
					i	QLMRELDQRTEDKKAEIDILAAEYISTVKTLS
1		ļ		ľ		PDQRVERLQKIQNAYSKCKEYSDDKVQLAM
1		i	i		1	QTYEMVDKHIRRLDADLARFEADLKDKMEG
1					İ	SDFESSGGRGLKKGRGQKEKRGSRGRGRRTS
						EEDTPKKKKHKGG
925	2275	A	7994	447	589	LPCSFCAQCMSSFERVWLQQSHFHNPRWNSR
				'''	1 505	SPIRCYCQHWPHCVHC
926	2276	A	7996	925	600	OPONIOCENIA DE CONSTRUCTOR DE LOS LOS DELOS
920	22/0	^	1990	925	582	GPCKVCCITLAIMLQCHSFYRKDVQVEHPKS
			1		1	LNPKYSQIENFLSADMALKRKCLLSISDLDFW
1						IWDAQPVGIMQTLQNLKKIPNPGCFWSQAFQI
		L				RDTQPILPLGGRYYTTIRQ
927	2277	Α	7998	2	353	RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV
		l			i	CLLLVTLALCCYQANAEFCPALVSELLDFFFI
					}	SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ
						MSLQKRSLIAEVLVKILKKCSV
928	2278	A	8004	120	500	
920	22/8	/A.	8004	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAG
						ARGLRATYHRLLDKVELMLPEKLRPLYNHPA
						GPRTVFFWAPIMKWGLVCAGLADMARPAEK
						LSTAQSAVLMATGFIWSRYSLVIIPKNWSLFA
						VNFFVGAAGASQLFRIWRYNQELKAKAHK
929	2279	Ā	8007	2	1016	EFARRVFIAAREMSLLRSLRVFLVARTGSYP
				_		AGSLLRQSPQPRHTFYAGPRLSASASSKELLM
<b>1</b>						KLRRKTGYSFVNCKKALETCGGDLKQAEIWL
			l	•		
] ]						HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE
		•				GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL
						GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA
						GPDREGSLKDQLALAIGKLGENMILKRAAWV
1 1						KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG
J J						ALVICETSEQKTNLEDVGRRLGQHVVGMAPL
[					:	SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ
	İ					YVQPQGVSVVDFVRFECGEGEEAAETE
930	2280	A	8008	3	1679	
550	2200	Λ.	avv <b>o</b>	,	10/7	NSRVWGPWTEPSAGSLRPMARKQNRNSKEL
1 1			}			GLVPLTDDTSHAGPPGPGRALLECDHLRSGV
					·	PGGRRKDWSCSLLVASLAGAFGSSFLYGYN
[					·	LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL
					1	TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK
1				-		HTLLANNGFAISAALLMACSLQAGAFEMLIV
					·	GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG
			1			QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF
1						
] ]						GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA
1		i				RAVKAFQTFLGKADVSQEVEEVLAESRVQRS
1 1						IRLVSVLELLRAPYVRWQVVTVIVTMACYQL
) l		l		1		CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG
l l					l	IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF
ı İ			1			GTLTITLTLQDHAPWVPYLSIVGILAHASFCSG
1 1			j			PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN
1 1	l	!	ļ	ļ	. 1	FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL
	•	1	ļ	ļ		VEW DETENDITY ACTOMATICAL AND DESCRIPTION OF A STREET AND DESCRIPTION ASSESSMENT OF A STREET AND DESCRIPTION ASSESSMENT OF A STREET AND DESCRIPTION ASSESSMENT OF A STREET ASSESSMENT O
, 1		- 1	1			YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI
L	2201	<u> </u>	0006		200	DSAVTDGKINGRP
931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL
1 1	1	·	1			NRNARRKAAPRIECSHIRHAWDHAKSVRQNL
[l			- 1			AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ŀ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	5-4	/=possible nucleotide deletion, \=possible
1		ŀ		sequence		nucleotide insertion
		<b></b>		Soquotico		PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
		l	İ		1	LIDYVRYMVENHGEDYKAMARDEKNYYQD
l i		ł				TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
						MEVE TRAFTFAEWQDFLDSDQRRR
932	2282	A	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
332	2202	1 ~	.0011	412	1	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
ĺ		ĺ			ĺ	, , , ,
						QGRGQIPIPCPWPPPPPPPPPPGSPGCRQFHQ
					ľ	SLEAKARHPASVREMRGKVKMRRALRRAPA
022	2002		0010	147	1077	STRASSRQPNPK
933	2283	Α	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
						ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
			1			NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
f 1	!		<u> </u>		1	RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
	ı	1		•		PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
						NASQLITQRAQVSLLIRRELTERAKDFSLILDD
						VAITELSFSREYTAAVEAKQVAQQEAQRAQF
		ł				LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
1 1		i			ł	SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
ii						DNLVLNLQDESFTRGSDSLIKGKK
934 ·	2284	Α	8023	255.	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
				•	i	RLRKFRELHLMRNEARKLNHQEVVEEDKRL
	ļ		1		İ	KLPANWEAKKARLEWELKEEEKKKECAARG
						EDYEKVKLLEISAEDAERWERKKKRKNPDLG
						FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
						KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
i l						KQIEKRDKYSRRRPYNDDADIDYINERNAKF
						NKKAERFYGKYTAEIKQNLERGTAV
935	2285	Α	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
						QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL
i				Maria San		SQHSSPAPMYSQTFHILVLG
936	2286	A	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
177		••	0000		,	FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLO
1 1						IWDTAGOERFRTITTAYYRGAMGIMLVYDIT
1 . 1						NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
			•			DVNDKRQVSKERGEKLALDYGIKFMETSAK
]						ANINVENAFFTLARDIKAKMDKKLEGNSPQG
				*		SNQGVKITPDQQKRSSFFRCVLL
937	2287	<u> </u>	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
937		<u>A</u>				
ا قدر ا	2288	Α	8052	675	-1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
]						PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
,						ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
						LREYQTRQDQCIYNTTYLNVQRENGTISRYV
						GGQEHFAHLLILRDTKTYMLAFDVNDEKNW
						GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
						VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	A	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
	+					AEQLKWSAELARLGESIMDGKQGGMDGSKP
}						AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
.						IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
.						CLGGHLSCVKILLKHGAQVNGVTADWHTPL
i i				·		FNACVSGSWDCVNLLLQHGASVQPESDLASP
			ĺ	Í		IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
			,	ļ		PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTOAKN
						AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
						CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
				l		L
940	2290	A	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
' '		**	3030	-		ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI
<u></u>			l			THE THE PLOT ON SON ALCHOOLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible mucleotide insertion  VVDTVMCPNMPNKSVLLYTLSFTYIFIFVIGMI
-	•					ANSVVVWVNIQAKTTGYDTHCYILNLAIADL WVVLTIPVWVVSLVQHNQWPMGELTCKVTH LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS RKKMVRRVVCILVWLLAFCVSLPDTYYLKT VTSASNNETYCRSFYPEHSIKEWLIGMELVSV VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI PFTCRLEHALFTALHVTQCLSLVHCCVNPVL YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA SRVSETEYSALEQSTK
941	2291	A	8059	73		DMAGLMTIVTSLLFLGVCAHHIPTGSVVLPS PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT MMFWYRQQPGQSLTLLATANQGSEATYESGF VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA VFEPSEAEISHTQKATLVCLATGFYPDHVELS WWVNGKEVHSGVSTDPQPLKEQPALNDSRY CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE NDEWTQDRAKPVTQIVSAEAWGRADCGFTS ESYQQGVLSATILYEILLGKATLYAVLVSALV LMAMVKRKDSRG
943	2293	A	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI TASERLRRPRATARLRAHAAPPEPPLAVFAP PSDRKELLALPVACDPVIASVMSWVQAASLI QGPGDKGDVFDEEADESLLAQREWQSNMQR RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA EVILNYGRLRGTLSALLSWCHLHNNNSTLINK INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL DSIEDMDLCHVVPAEKKIDEAKDERLCENNA EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK PKPHMDFGTDSQF
944	2294	Α	8073	1	797	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN DPDIEMA WAMRAMQHAEVYYKLISSVDPQF LKLTKVDDQIYSEFKNFETLRIDVLDPEELK SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD CSQGYTEENTIFAPRIQFFAIEIARNREGYNKA VYISVQDKEGEKGVNNGGEKRADSGEEENT KNGGEKGADSGEEKEEGINREDKTDKGGEK GKEADKEINKSGEKAM
945	2295	A	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL SADRRVLGLREWGRPASERECSLCQRLKREL NMGDVEKGKKIFIMKCSQCHTVEKGGKHKT GPNLHGLFGRKTGQAPGYSYTAANKNKGIIW GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER ADLIAYLKKATNE
946	2296	A	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF VAIFAVPLILGQEYEDEERLGEDEYYQVVYY YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK DITEAIETTISLETARADHPKPVTVKPVTTEPQ SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC KKVGRRLLMTLWMGVWQEEIGR
947	2297	Α .	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF SRARAPAHSLRAALSLASSARSWGAVSRDRG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion PCPPAIMYQSSNKC
948	2298	В	8093	3905	846	MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSALAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGGGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEGQGSPSEMVHSSEFRVDNNHLLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKQFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTTTKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSP FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYFSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKRHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHVYDLPVEGNEPPASFHKDTNALPPTVFKD NPMSTPKDSNPKA

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted begiming nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
950	2300	A	8100			MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLNIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFIYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES
956	2306	A	8157	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI
						QCQNKGWDGYDVQWECKTDLDIAYKFGKT VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL GLQKLKESGKQHGFASFSDYYYKWSSADSC NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ NTGHGATSGFGSAFTGQQGYENSGPGFWTGL GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA SGYGGTRRR
957	2307	A	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVQ DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA HNEVGTAQCNVTLEVSTOPGAAVVAGAVVG TLVGLGLLAGLVLLYHRRGKALEEPANDIKE DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS QAGSLV
958	2308	A	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP LNATLVITFEITFRSKNITILELPDEVVVPPGVT NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	Α	8163	521	1345	GERAGRRRGRLGVWAQPQPLLPRPVGSRRE MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI ASPDVKLNLGGDFIKESTATTFLRQRGYGWL LEVEDDDPEDNKPLLEELDDLKDIYYKIRCV LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK KPLLIYPIFLLYIYFLSLYTGV
960		A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN LYSQLNALQFTVDERSILWLNQFLLDLKQSL NQFMAVYKLNDNSKSDEHVDVRVDGLMLK FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI HLLVHVHKHVSMQINHYQYLLLLFLHESLILL SENLRKDVEAVTGSPASQTSICIGILLRSAELA LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM SVDLSHIPLKDPLLFKSASDTNLQKGISFMDY LSDKHLGKISEDESSGLVYKSGSGEIGSETSD KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI LSSTLTSKGNETIESIFKAEDLLPEAASLSENL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	ncuce		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP
		ł				LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL
1		}	}			ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN
					'	AGANLQNYGETSPDAISTNSEGAQENHDDLM
						SVVVFKITGVNGEIDIRGEDTEICLQVNQVTP DOLGNISLRHYLCNRPVGSDOKAVIHSKSSPE
			ľ			ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
						EFLTSSLMNIQHFLEDETVATVMPMKIQVSNT
						KINLKDDSPRSSTVSLEPAPVTVHIDHLVVER
-						SDDGSFHIRDSHMLNTGNDLKENVKSDSVLL
			ļ		l	TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN
						FPEFSFDFTREQLMEENESLKQELAKAKMAL AEAHLEKDALLHHIKKMTVE
961	2311	Α	8172	1442	682	TAAMSIFTPTNOIRLTNVAVVRMKRAGKRFEI
						ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN
			ļ			VSKGQVAKKEDLISAFGTDDQTEICKQILTKG
1						EVQVSDKERHTQLEQMFRDIATIVADKCVNP
						ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL
1					ļ	KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI
						DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS
					!	VLRRMQKKYWKTKQVFIKATGKKEDEHLVA
						SDAELDAKLEVFHSVQETCTELLKIEKYQLR
963	2313	A	8181	13	2215	LNGMKS AEGCAERRGTEPVVELSMSWESGAGPGLGSQ
703	2313	,	0101	13	2213	GMDLVWSAWYGKCVKGKGSLPLSAHGIVV
						AWLSRAEWDQVTVYLFCDDHKLQRYALNRI
						TVWRSRSGNELPLAVASTADLIRCKLLDVTG
						GLGTDELRLLYGMALVRFVNLISERKTKFAK
						VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET
					•	WELEEFREGIEEEDQEEDKNIVVDDITEQKPE
						PODDGKSTESDVKADGDSKGSEEVDSHCKK
						ALSHKELYERARELLVSYEEEQFTVLEKFRYL
						PKAIKAWNNPSPRVECVLAELKGVTCENREA
						VLDAFLDDGFLVPTFEQLAALQIEYEENVDL NDVLVPKPFSQFWOPLLRGLHSONFTOALLE
ľ						RMLSELPALGISGIRPTYILRWTVELIVANTKT
				İ	!	GRNARRFSAGQWEARRGWRLFNCSASLDWP
1						RMVESCLGSPCWASPQLLRIIFKAMGQGLPD
						EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK
						SPYTLDSLYWSVKPASSSFGSEAKAQQEEQ
					•	GSVNDVKÉEEKEEKEVLPDQVEEEEENDDQE EEEEDEDDEDDEEEDRMEVGPFSTGOESPTA
						ENARLLAQKRGALQGSAWQVSSEDVRWDTF
						PLGRMPGQTEDPAELMLENYDTMYLLDQPV
						LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS
0.00	001:		64.0		1505	NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRRNFRDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLOLPAGPFSRCRWDPVSSPR
'						PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP
			, i			NVGKSTFFNVLTNSQASAENFPFCTIDPNESR
						VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG
						LVKGAHNGQGLGNAFLSHISACDGIFHLTRA
-[						FEDDDITHVEGSVDPIRDIEIIHEELQLKDEEMI
						GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS
					· '	WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD
		.	ĺ			KYDPGALVIPFSGALELKLOELSAEEROKYLE

SEQ ID	OFO ID	N/ot	LCCO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Lunc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	}	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		peptide		/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
				· · · · · · · · · · · · · · · · · · ·		ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV
		1				RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV
		1	ĺ	f .		MKYEDFKEEGSENAVKAAGKYRQQGRNYIV
						EDGDIIFFKFNTPQQPKKK
965	2315	Α	8195	1437	594	RSFSLSFSLLSPSEMMALGAAGATRVFVAMV
		ļ			}	AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL
						GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY
		1				QPYPCAEDEECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS
		ĺ	Í		ĺ	DQNHFRGEIEETITESFGNDHSTLDGYSRRTT
						LSSKMYHTKGQEGSVCLRSSDCASGLCCARH
		Ì	İ			FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ
						RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH
966	2316	A	8207	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW
						SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI
			1			DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY
			1		ĺ	WVDLERQLLQRVFLNGSRQERVCNIEKNVSG
						MAINWINEEVIWSNQQEGIITVTDMKGNNSHI
			ļ			LLSALKYPANVAVDPVERFIFWSSEVAGSLY
		Ì				RADLDGVGVKALLETSEKITAVSLDVLDKRL
		·				FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ
			ĺ			HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG
						KDMVRINLHSSFVPLGELKVVHPLAQPKAED
					i	DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC
	•	١.	1		!	MCAEGYALSRDRKYCEGNDWKYCEDVNEC   AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD
						GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF
						CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL
	· ·		1			SPVSWECDCFPGYDLQLDEKSCAASGPQPFL
			<b>(</b>			LFANSQDIRHMHFDGTDYGTLLSQQMGMVY
						ALDHDPVENKIYFAHTALKWIERANMDGSQ
						RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK
			1			SLIGRSDLNGKRSKIITIENISOPRGIAVHPMAK
						RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW
						PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK
			1			RRRLTQNDVGHPFAVAVFEDYVWFSDWAMP
ĺ						SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP
						LAKPGADPCLYQNGGCEHICKKRLGTAWCS
			Ì			CREGFMKASDGKTCLALDGHQLLAGGEVDL
			<b>i</b>			KNQVTPLDILSKTRVSEDNITESQHMLVAEIM
.			]			VSDQDDCAPVGCSMYARCISEGEDATCQCLK
						GFAGDGKLCSDIDECEMGVPVCPPASSKCINT
						EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS
1		İ '		;		CGENASCTNTEGGYTCMCAGRLSEPGLICPD
ł						STPPPHLREDDHHYSVRNSDSECPLSHDGYCL
						HDGVCMYIEALDKYACNCVVGYIGERCQYR
						DLKWWELRHAGHGQQQKVIVVAVCVVVLV
						MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS
[						RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD
						LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ
						LCGMGTEQGCWIPVSSDKGSCPQVMERSFH
						MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL
						DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL
					'	RLHHRFRALDRNKKGYLSRMDLQQIGALAV
1						NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP
						VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY
						DLDRDGKISRHEMLQVLRLMVGVQVTEEQL
		i				
						ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	пос	in NO.	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	Seq-	ł	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1	1	peptide		/=possible nucleotide deletion, \=possible
0/0	2210		0011	sequence	465	nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT
1	,	}				YMDNWRQNTTAEQEALQAKVDAENFYYVIL YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP
						YHQYIVEDWQEKYKSQILNLEESKATIHENIG
						AAGFKMSP
969	2319	Α	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPGOAPRWSRWR
			02.0	•	1,550	VPGRLLLLLPALCCLPGAARAAAAAAGAGN
						RAAVAVAVARADEAEAPFAGONWLKSYGY
						LLPYDSRASALHSAKALQSAVSTMQQFYGIP
	[					VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR
	İ			-		NKRYALTGQKWRQKHITYSIHNYTPKVGELD
						TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR
1 .	}					KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF
						PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL
						VAVHELGHALGLEHSSDPSAIMAPFYQYMET
						HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL PVRRIHSPSERKHEROPRPPRPPLGDRPSTPGT
				8		KPNICDGNFNTVALFRGEMFVFKDRWFWRL
	1					RNNRVQEGYPMQIEQFWKGLPARIDAAYER
	ļ					ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG
	ļ					ELGSCLPREGIDTALRWEPVGKTYFFKGERY
	1					WRYSEERRATDPGYPKPITVWKGIPQAPQGA
						FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY
						PRNILRDWMGCNQKEVERRKERRLPQDDVDI
						MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI
970	2200	_	0016	1005	2000	FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN
						DSLRTNVFVRFQPETIACACIYLAARALQIPLP TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN
				-		YELLEKEVEKRKVALQEAKLKAKGLNPDGTP
1						ALSTLGGFSPASKPSSPREVKAEEKSPISINVK
						TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR
						SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY
						SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR
						DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA
						KKHRHERGHHRDRRERSRSFERSHKSKHHGG
						SRSGHGRHRR
971	2321	A	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE
, ,						RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN
		1				NKLANYTNLSOGVVEHEEDEESRRREAKAPR
						MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG
						VLESFLIVAMCCTCTMLTAISMSAIATNGVVP
1 1		ł	ł	ł		AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA
1		-		.		GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA
] ]		j	j	~ }		AMLHNMRVYGTCTLVLMALVVFVGVKYVN
					ļ	KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV.
						CLLGNRTLSRRSFDACVKAYGIHNNSATSAL
			-	-		WGLFCNGSQPSAACDEYFIQNNVTEIQGIPGA
		1				ASGVFLENLWSTYAHAGAFVEKKGVPSVPV
		l				AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY
	,			.		LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM
1		ĺ	[			LAWPSPWVIVIGSFFSTCGAGLOTLTGAPRLL
			'			QAIARDGIVPFLQVFGHGKANGEPTWALLLT
1		l				VLICETGILIASLDSVAPILSMFFLMCYLFVNL
] ]		J	ļ			ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL
				}		CLALMFICSWYYALSAMLIAGCIYKYIEYRG
		1		Į		AEKEWGDGIRGLSLNAARYALLRVEHGPPHT KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMA WPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVILVRGGGREVTITYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI
973	2323	A	8237	279	4610	LERKAKSRQVGKEKGKYKEELIEKMQE GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEEGAGGRQDPSRRSIRLR PLPSPSPSAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRGGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRTTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGQTPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPSS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGSS GGLGPPGRPYGAIPGQHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL G

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	ng to first amino acid residue of peptide	to last amino acid residue of peptide sequence	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, \=possible
		,		sequence		nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA P
976	2326	A .	8257	298	7086	GNMACWPOLRLLLWKNLTFRRRQTCOLLLE VAWPLFIFLLISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTTMENSQEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWPI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTIILSTHHMDEADVL GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT LVKKDVESSLSSCRNSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVSAISNLIRKHVSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID
						DRLSDLGISSYGISETTLEEIFLKVAEESGVDA ETSDGTLPARRNRRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l-Isoleucine, K-Lysine, L-Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	]	1	914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	Soquettee	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				-		LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY
						NEQYTFVSNDAPEDTGTLELLNALTKDPGFG
						TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM
1						DLFQNGNWTMQNPSPACQCSSDKIKKMLPV
	ŀ					CPPGAGGLPPPQRKQNTADILQDLTGRNISDY   LVKTYVOIIAKSLKNKIWVNEFRYGGFSLGVS
						NTQALPPSQEVNDATKQMKKHLKLAKDSSA
						DRFLNSLGRFMTGLDTRNNVKVWFNNKGW
						HAISSFLNVINNAILRANLQKGENPSHYGITAF
						NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA
						MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI
						YWLSNFVWDMCNYVVPATLVIIIFICFQQKSY
						VSSTNLPVLALLLLLYGWSITPLMYPASFVFK   IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN
1						KLNNINDILKSVFLIFPHFCLGRGLIDMVKNO
						AMADALERFGENRFVSPLSWDLVGRNLFAM
				4.		AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN
						DEDEDVRRERQRILDGGGQNDILEIKELTKIY
1						RRKRKPAVDRICVGIPPGECFGLLGVNGAGK
	!					SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV
					-	HQNMGYCPQFDAITELLTGREHVEFFALLRG
1						VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD
[			->-			PKARRFLWNCALSVVKEGRSVVLTSHSMEEC
						EALCTRMAIMVNGRFRCLGSVQHLKNRFGD
						GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK
1						EKHRNMLQYQLPSSLSSLARIFSILSQSKKRLH
					•	IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL
977	2327	Α	9260	3	1567	SLHKNQTVVDVAVLTSFLQDEKVKESYV
711	2321	А	8260	3	1567	IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV
] .						TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM
•			8	•		SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH
1.						YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS
						LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM
				j		VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV
						MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG
						MYAYAGWFYLNFVTEEVENPEKTIPLAICISM   AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT
						PSERLLGNPSLAVPIFVALSCFGSMNGGVFAV
	1	ł	}			SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV
			Ì			LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA
				l		VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC
]					ļ	LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII
070	2220	<del></del>	P261	2	2165	WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
978	2328	A	8261	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP
			1			LADAASMSGVRAVRISIESACEKQVHEVGLD
				ļ		GTETYLPPLSMSONLARLAORIDFSOGSGSEE
1 1	ľ	ľ	1	ľ		EEAAGTEGDAQEWPGAGSSADQDDEEGVVK
			1			FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI
		ļ	1			VRDKKFMTLDPVSQDALPPKQNPQTLQLISK
<u> </u>			İ			KKSLAGAAQILLKGAERLTKSVTENQENKLQ
1			ľ			RDFNSELLRLRQHWKLRKVGDKILGDLSYRS
!	j	1				AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN
			ļ			LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI
			ļ	}		FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ
	ſ	ĺ	-	İ	[	LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE
L						HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR
					-	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valline, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND
					•	VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG NASAITVASPSGDYAISVRNGPESGSKIMVQF PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ WNKMEGRNFVYKMELLMSALSPCLL
979	2329	A	8289	2	1053	FVWNPRGGRKRRQAAVTQAATRASGTPSP RDGTMTQGKLSVANKAPGTEGQQQVHGEKK EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV ALFTFCDPVKDYVQANPGWYWASYAVFFAT YLTLACCSGPRHFPWNLILLTVFTLSMAYLT GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ TKFDFTSCQGVLPVLLMTLFFSGLILAILLPFQ YVPWLHAVYAALGAGVFTLFLALDTQLLMG NRRHSLSPEEYIFGALNIYLDIIYIFTFFLQLFG TNRE
980	2330	A	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY MNAAMVHINRALKLIIRLFLVEDLVDSLKLA VFMWLMTYVGAVFNGITLLILAELLIFSVPIV YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG IAKKKAE
981	2331	A	8308	186	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD YDLCASCYESGATTTRHTTDHPMQCILTRVD FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET SLQEHVTSEHAETSTEVICPICAALPGGDPNH VTDDFAAHLTLEHRAPRDLDESSGVRHVRR MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ LNSSGPSASQLQQLQMQLQLERQHAQAARQ QLETARNATRRTNTSSVTTITGSTATTNIAN TESSQQTLQNSQFLLTRLNDPKMSETERQSM ESERADRSLFVQELLLSTLVREESSSDEDDR GEMADFGAMGCVDIMPLDVALENLNLKESN KGNEPPPPPL
982	2332	A	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV AAALLVGFILFLTRSRGRAASAGQEPLHNEEL AGAGRVAQPGPLEPEEPRAGGRPRRRDLGS RLQAQRRAQRVAWAEADENEEEAVILAQEE EGVEKPAETHLSGKIGAKKLRKLEEKQARKA QREAEEAEREERKRLESQREAEWKKEEERLR LEEEQKEEEERKAREEQAQREHEEYLKLKEA FVVEEEGVGETMTEEQSQSFLTEFINYIKQSK VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT GVIDDRGKFTYITPEELAAVANFIRQRGRVSIA ELAQASNSLIAWGRESPAQAPA
983	2333	A .	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP CPQPLRSPSLDNPTPFPNLGPSENPLKRLLVPG EEWEFEVTAFYRGRQVFQQTISCPEGLRLVGS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPFIVGSLGPPDLITFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY
984	2334	A	8321	1	1243	KAYLQDLVEGMDFQGPGES  ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATRRILAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSLSATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYANPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89		KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADTPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A	8335	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RTRIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	A	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL KSLHPMS
990	2340	A	8361	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH QFKLCIMRRSKGRAEKS  SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL
						EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE DATLENGCLWFIPGSHTSGVSRRMVRAPVGS APGTSFLGSEPARDNSLFVPTPVQRGALVLIH GEVVHKSKQNLSDRSRQAYTFHLMEASGTT WSPENWLQPTAELPFPQLYT
992	2342	Α .	8370	906		MALSGNCSRYYPREQGSAVPNSFPEVVELNV GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD RKWGFITVGYRGSCTLGREGQADAKFRRVPR ILVCGRISLAKEVFGETLNESRDPDRAPERYTS RFYLKFKHLMGAPASNFILGFWGLGQNQDK HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA QEVSRRRRWLGDPEHL
993	2343	A	8379		2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE DMVTTASTYLFEATEKRFFFKNVSILIPENWK ENPQYKRPKHENHKHADVIVAPPTLPGRDEP YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST WEVISNSEDFKNTIIPMVTPPPPPVFSLLKIRQRI VCLVLDKSGSMGGKDRLNRMQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI VHFIALGRAADEAVIEMSKITGGSHFYVSDEA QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP SISLWDPSGTIMENFTVDATSKMAYLSIPGTA KVGTWAYNLQAKANPETLTITVTSRAANSSV PPITVNAKMNKDVNSFPSPMIVYAEILQGYVP VLGANVTAFIESQNGHTEVLELLDNGAGADS FKNDGVYSRYFTAYTENGRYSLKVRAHGGA NTARLKLRPPLNRAAYIPGWVNGEIEANPP RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL PLPDQYPPSQITDLDATVHEDKIILTWTAPGD NFDVGKVQRYIIRISASILDLRDSFDDALQVN TTDLSPKEANSKESFAFKPENISEENATHIFIAI KSIDKSNLTSKVSNIAQVILFIPQANPDDIDPT PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG TQDEILKMRNTFAELKNSLEALSSRMDQAEE RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
995	2345	A	8390	194	3421	PSSWDYRACLS AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			Ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
1 1				sequence		nucleotide insertion
						DFLSMKQSPALAPEERCRRAGSPKPVLRADD
			•			NNMGNGCSQKLATANLLRFLLLVLIPCICALV
		ļ ·	ł			LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI
1 1		1	Ì			QGSDVILTNTTYNQSTVVSTAHPDQHVPAWT
1 1			1			TDASLPGDQSHRNTSACMNTTHSQCQMLPYH
1 1		1	1			ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY
						QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE
	•					AAKEGCESVLGMVNYSWPDFLRCSQFRNQT
1 1						ESSNVSRICFSPQQENGKQLLCGRGENFLCAS
1 1						GICIPGKLQCNGYNDCDDWSDEAHCNCSENL
1 1						FHCHTGKCLNYSLVCDGYDDCGDLSDEONC
1 1						DCNPTTEHRCGDGRCIAMEWVCDGDHDCVD
1 1			i I			KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG
i !					Ι,	DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP
l l						
					. 1	CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM
1 1						NLPYNSTSYPNYFGHRTQKEASISWESSLFPA
						LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP
						CRALCEHSKERCESVLGIVGLQWPEDTDCSQ
						FPEENSDNQTCLMPDEYVEECSPSHFKCRSGQ
						CVLASRRCDGQADCDDDSDEENCGCKERDL
1 1		ĺ				WECPSNKQCLKHTVICDGFPDCPDYMDEKN
}				i		CSFCQDDELECANHACVSRDLWCDGEADCS
1						DSSDEWDCVTLSINVNSSSFLMVHRAATEHH
						VCADGWQEILSQLACKQMGLGEPSVTKLIQE
						QEKEPRWLTLHSNWESLNGTTLHELLVNGQS
						CESRSKISLLCTKQDCGRRPAARMNKRILGGR
						TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW
						VLTVAHCFEGRENAAVWKVVLGINNLDHPS
1 1						VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
						DISETGYVRPVCLPNPEQWLEPDTYCYITGW
1 1					, i	GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK
						TITTRMICAGYESGTVDSCMGDSGGPLVCEK
						PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS
						YFVEWIKRQIYIQTFLLN
996	2346	A	8392	199	3085	KVILSSEMSKTNKSKSGSRSSRSRSRSRSRSRS
""	25.0	1.	0372	***	5005	FSKSRSRSRSLSRSRKRRLSSRSRSRSYSPAHN
						RERNHPRVYONRDFRGHNRGYRRPYYFRGR
1						NRGFYPWGQYNRGGYGNYRSNWQNYRQAY
						SPRRGRSRSRSPKRRSPSPRSRSHSRNSDKSSS
1 1						DRSRRSSSSRSSSNHSRVESSKRKSAKEKKSSS
1						KDSRPSQAAGDNQGDEVKEQTFSGGTSQDTK
1						ASESSKPWPDATYGTGSASRASAVSELSPRER
1						
1						SPALKSPLQSVVVRRRSPRPSPVPKPSPPLSST
1 1						SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
1 1						SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
		0				YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE
						KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
1						SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
}						DKAKGRKESEFDDEPKFMSKVIGANKNQEEE
] [						KSGKWEGLVYAPPGKEKQRKTEELEEESFPE
1						RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
	i			]		AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS
						SFSITREAQVNVRMDSFDEDLARPSGLLAQER
] !						KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
1				İ		SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
					(A)	KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
		į				HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
			٠	0.03	n e	HGLAHDEMKSPREPGYKAEGKYKDDPVDLR LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
			٠			HGLAHDEMKSPREPGYKAEGKYKDDPVDLR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	sed-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		ŀ	' ' '	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	]	]	]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		ł		sequence		nucleotide insertion
						GTKDFVGPSERGGGRARGTFQFRARGRGWG
ł	ì	l			ĺ	RGNYSGNNNNNSNNDFQKRNREEEWDPEYT
		l		•		PKSKKYYLHDDREGEGSDKWVSRGRGRGAF
		ļ	İ			PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE
						DDESGTENREEKDNIQPTTE
997	2347	Α	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS
		ł				KQENLESLEATGREEEGGQGPPVTTKGVLLA
						LLMAGLALQPGTALLCYSCKAQVSNEDCLQ
					L	VENCTQLGEQCWTARIREWGDDSRQA
998	2348	Α	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN
		]				MSDPRRPNKVLRYKPPPSECNPALDDPTPDY
	١.	İ				MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI
i		<u> </u>				SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ
	20.15		0.40			NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL
1				l .		LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL
1						VSKKANQQLNFTEAKEACRLLGLSLAGKDQ
				•		VETALKASFETCSYGWVGDGFVVISRISPNPK
						CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW
						TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS
						VASPYSTIPAPITTPPAPASTSIPRRKKLICVTE
	, i					VFMETSTMSTETEPFVENKAAFKNEAAGFGG
						VPTALLVLALLFFGAAAGLGFCYVKRYVKAF   PFTNKNQQKEMIETKVVKEEKANDSNPNEES
						KKTDKNPEESKSPSKTTMRCLEAEV
1000	2350	A	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT
1			0.00	_	'''	AAIFTADGNMISASTLMDILLMNDFKLVINKI
						AYDVQCPKREKPSNEHTAEMEHMKSLVHRL
						FTILHLEESQKKREHHLLEKIDHLKEQLOPLE
						QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA
						LAWLTWWVYSWDIMEPVTYFITFANSMVFF
						AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ
}						HFDVQQYNKLKEDLAKAKESLKQARHSLCL
L :						QMQVEELNEKN
1001	2351	Α	8410	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARG
						TGALLLRGSLLASGRAPRRASSGLPRNTVVLF
						VPQQEAWVVERMGRFHRILEPGLNILIPVLDR
			1	ĺ		IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV
						LYLRIMDPYKASYGVEDPEYAVTQLAQTTM
						RSELGKLSLDKVFRERESLNASIVDAINQAAD
}						CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE
	0					RRKRATVLESEGTRESAINVAEGKKQAQILAS
					.•	EAEKAEQINQAAGEASAVLAKAKAKAEAIRI
, ·						LAAALTQHNGDAAASLTVAEQYVSAFSKLA
			ĺ	100		KDSNTILLPSNPGDVTSMVAQAMGVYGALT
		- 1		ſ	İ	KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR
1002	2252		DA21	124	041	VKMS
1002	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY
	1			Ì		YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG
.						NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL
						AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC
		- 1			ļ	PKCTQIINGSVDVDTEDRQKRKPESDGRTAK
		1				ALRSLQFTNPGRQTEFAPETGKREKRRLTKN
		İ	]			ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK
·		l	,			WLYEQIEIEGGEIIHNKHAG
1003	2353	$\overline{\mathbf{A}}$	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR
	ددب	**	5721	·	1710	CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT
		,	1		į	MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SLDTENIDEILNNADVALVNFYADWCRFSQM LHPIFEEASDVIKEEFPNENQVVFARVDCDQH SDIAQRYRISKYFTLKLFRNGMMMKREYRGQ RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
•						KRNIIGYFEQKDSDNYRVFERVANILHDDCAF LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY LGAMTNFDVTYNWIQDKCVPLVREITFENGE ELTEEGLPFLILFHMKEDTESLEIFQNEVARQL ISEKGTINFLHADCDKFRHPLLHIQKTPADCP VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDL HSGKLHREFHHGPDPTDTAPGEQAQDVASSP PESSFQKLAPSEYRYTLLRDRDEL
1004	2354.	A	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM ACAAARSPADQDRFICIYPAYLNNKKTIAEGR RIPISKAVENPTATEIQDVCSAVGLNVFLEKN KMYSREWNRDVQYRGRVRVQLKQEDGSLC LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA DQSLQQGEGSKKGKGKKKK
1005	2355	A	8453	90	530	QSHETKMQSGTHWRVLGLCLLSVGVWGQD GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF SELEQSGYYVCYPRGSKPEDANFYLYLRARG NPGLQNRYHRLFREDHSKGHSQ
1006	2356	A	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC QLCIFN
1007	2357	A	8459	43	553	GAGAGGDWAAMDKLKKVLSGODTEDRSGL . SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA FWWHNKGLALIFCILQSLALTWYSLSFIPFAR DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS DPRWGCVGPSMPTSTCLPGAVEASTTKASLP KCPVDSSLPTPEACFL
1009	2359	A .	8465	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP NETIIVLPSNVINFSQAEKPEPTNQGQDSLKKH LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN FTQVTSTLLINSAYPFIGPFFFIISGSLSIATEKRL TKLLVHSSLVGSILSALSALVGFIILSVKQATL NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD CYTAKASLAGTLSLMLICTILLEFCLAVLTAVL RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT HDCGYEELLTS
1010	2360	A	8468	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA HRVALCHLAGCQEQAAWYHTLQILFFLVSAY FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICT LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	A	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT GTLETQFTCPFCNHEKSCDVKMDRARNTGVI SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM RMKYGGQEFWADLNAMNVYETTEFDQLRR LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SVIRLIEEANSRGLKEVRFMMWNNHYILHNS FFRREIKRRPLFRSCFILLPYLQTLGGVPTQAP PPLEATSSSQIICPDGVTSANFYPETWVYMHP SQDFIQVPVSAEDKSYRIIYNLFHKTVPEFKYR ILQILRVQNQFLWEKYKRKEYMNKMFGR DRIINERHLPHGTSQDVVDGICKHNFDPRVCG KHATMFGQGSYFAKKASYSHNFSKKSSKGV HFMFLAKVLTGRYTMGSHGMRRPPPVNPGS VTSDLYDSCVDNFFEPQIFVIFNDDQSYPYFVI QYEEVSNTVSI
1013	2363	A	8488	2	517	IENCRTRLRQAWHEVCGNKMAAPIPQGFSCL SRFLGWWFRQPVLVTQSAAIVPVRTKKRFTP PIYQPKFKTEKEFMQHARKAGLVIPPEKSDRS IHLACTAGIFDAYVPPEGDARISSLSKEGLIER TERMKKTMASQVSIRRIKDYDANFKIKDFPE KAKDIFIEGSPLY
1014	2364	A .	8501	363	17	YIRTGYVYICIIYAQLMYTYYIRTAYVYICILY AQLMYTYVLYTHSLCIHMYSIRTAYVYICIIY AQIMYTYVFYTHRLCIHMYSIRTDYVYICILY AQLMYTYVFYTHSYMSDE
	2365	A	8504	3	2190	NSSEHFSQAPQRLSFYSWYGSARLFRFRVPPD AVLLRWLLQVSRESGAACTDAEITVHFRSGA PPVINPLGTSFPDDTAVQPSFQVGVPLSTTPRS NASVNVSHPAPGDWFVAAHLPPSSQKIELKG LAPTCAYVFQPELLVTRVVEISIMEPDVPLPQ TLLSHPSYLKVFVPDYTRELLLELRDCVSNGS LGCPVRLTVGPVTLPSNFQKVLTCTGAPWPC RLLLPSPPWDRWLQVTAESLVGPLGTVAFSA VAALTACRPRSVTIQPLLQSSQNQSFNASSGL LSPSPDHQDLGRSGRVDRSPFCLTNYPVTRED MDVVSVHFQPLDRVSVRVCSDTPSVMRLRL NTGMDSGGSLTISLRANKTEMRNETVVVACV NAASPFLGFNTSLNCTTAFFQGYPLSLSAWSR RANLIIPYPETDNWYLSLQLMCPENAEDCEQ AVVHVETTLYLVPCLNDCGPYGQCLLLRRHS YLYASCSCKAGWRGWSCTDNSTAQTVAQQR AATLLLTLSNLMFLAPIAVSVRRFFLVEASVY AYTMFFSTFYHACDQPGEAVLCILSYDTLQY CDFLGSGAAIWVTILCMARLKTVLKYVLFLL GTLVIAMSLQLDRRGMWNMLGPCLFAFVIM ASMWAYRCGHRRQCYPTSWQRWAFYLLPG VSMASVGIAIYTSMMTSDNYYYTHSIWHILL AGSAALLLPPPDQPAEPWACSQKFPCHYQIC KNDREELYAVT
1016	2366	A	8511	1	453	KWYPSGPVRIPGRFYYKLPAGHRRCRMAPAK KGGEKKKGRSAINEVVTREYTINIHKRIHGVG FKKRAPRALKEIRKFAMKEMGTPDVRIDTRL NKAVWAKGIRNVPYRIRVRLSRKRNEDEDSP NKLYTLVTYVPVTTFKNLQTVNVDEN
	2367	A	8513	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINT LSAKWADNFMAEGCGGSKEHSFQHPFLQAV GMFLGEFSCLAAFYLLRCRAAGQSDSSVDPQ QPFNPLLFLPPALCDMTGTSLMYVALNMTSA SSFQMLRGAVIIFTGLFSVAFLGRRLVLSQWL GILATIAGLVVVGLADLLSKHDSQHKLSEVIT GDLLIIMAQIIVAIQMVLEEKFVYKHNVHPLR AVGTEGLFGFVILSLLLVPMYYPAGSFSGNP RGTLEDALDAFCQVGQQPLIAVALLGNISSIA FFNFAGISVTKELSATTRMVLDSLRTVVIWAL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  SLALGWEAFHALQILGFLILLIGTALYNGLHR PLLGRLSRGRPLAEESEQERLLGGTRTPINDA S
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA YQLYQDPRNVWGFLAATSVTFVGVMGMRS YYYGKFMPVGLIAGASLLMAAKVGVRMLM TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLAVLL AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT PTGTHDHYMLEWFLTDRSGARPRLASAEMQ GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA QVGDERDYVCVVRAGAAGTAEAAARLNVF AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN GNPAPKITWYRNGQRLEVPVEMNPEGYMTS RTVREASGLLSLTSTLYLRLKDDRDASFHC AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG QSGTYGCRVEDYDAADDVQLSKTLELRVAY LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP KADGSWREGDEVTLICSARGHPDPKLSWSQL GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV MAVAVSVGLLLLVVAVFYCVRKKGGPCCRQ RREKGAP
1020	2370	A	8530	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG VLFSTGLLGNLLALGLLARSGLGWCSRRPLR PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL SSTLQLLAMALECWLSLGHPFFYRRHITLRLG ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV LATVLCNLGAMRNLYAMHRLQRHPRSCTR DCAEPRADGREASPQPLEELDHLLLALMTV LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI FIRPLRYRSRCSNSTNMESSL
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR KHDSGAADLERVTDYAEEKEIQSSNLETAMS
1022	2372	A	8537	94	541	VIGDRRSREQKAKQER RKERRRRRRMEAVVFVFSLLDCCALIFLSV YFIITLSDLECDYINARSCCSKLNKWVIPELIG HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS TAEALOL
1024	2374	A	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD FDWREVEILMFLSAIVMMKNRRSITVEQHIGN IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TCKPPLYMGPEYIKYFNDKTIDEELERDKRVT WIVEFFANWSNDCQSFAPIYADLSLKYNCTG LNFGKVDVGRYTDVSTRYKVSTSPLTKQLPT LILFQGGKEAMRRPQIDKKGRAVSWTFSEEN VIREFNLNELYQRAKKLSKAGDNIPEEQPVAS TPTTVSDGENKKDK
1025	2375	A	8546	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPA CRVPYCSVVCFRKHKEQCNPETRPVEKKIRS ALPTKTVKPVENKDDDDSIADFLNSDEEEDR VSLQNLKNLGESATLRSLLLNPHLRQLMVNL DQGEDKAKLMRAYMQEPLFVEFADCCLGIV EPSQNEES
1026	2376	Α	8547	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGS YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIVHISIFYPRVSLTDTGRFGVGMAIL SLLLKPLSCCFVYHMYRERGGELLVHTGFLG SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY
1027	2377	A	8557	1	340	DFLGPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYFEKAGLK
1028	2378	A	8569	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY
1029	2379	A	8572	1	578	AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV
1030	2380	A	8574	1352	372	DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF
1031	2381	A	8580	905	340	RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
1032	2382	A	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD VMNPSEILKGEKPQVRERGPYVYREFRHKSNI TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV MPNILVLGAAVMMENKPMTLKLIMTLAFTTL GERAFMNRTVGEIMWGYKDPLVNLINKYFP GMFPFKDKFGLFAELNNSDGLFTGFTGVQNI SRIHLVDKWNGLSKVDFWHSDQCNMINGTS GQMWPPFMTPESSLEFYSPEACRSMKLMYKE SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
					,	CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA ESGAMEGETLHTFYTQLVLMPKVMHYAQYV LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	A	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
1034	2384	A	8597	640	164	FCLLLSLVSSSLVSLSLCPPLTQA VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG GAGYVRSSQDLSCDFCNDVLARAKYLKRHG F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI HVYKKNGVGKVGDQILLAIKGQKKKALIVG HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI KTPIPTSLRKREGEYSKVLAIAQNFV
1036	2386	A	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV TQYLQPRSPEECKMFACAKLACTPSLIRAGSR VAYRPISASVLSRPEASRTGEGSTVFNGAQNG VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY AILGFALSEAMGLFCLMVAFLILFAM
1037	2387	A	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT GMVAHINNSRLKAKGVGQHDNAQNFGNQSF EELRAACLRKGELFEDPLFPAEPSSLGFKDLG PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL PTKNDKLVFVHSTERSEFWSALLEKAYAKLS GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP PQNLLRLLRKAVERSSLMGCSIEVTSDSELES MTDKMLVRGHAYSVTGLQDVHYRGKMETLI RVRNPWGRIEWNGAWSDSAREWEEVASDIQ MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV VVCTCLVALMQKNWRHARQQGAQLQTIGFV LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI FTNSREVSSQLRLPPGEYIIPSTFEPHRDADFL LRVFTEKHSESWELDEVNYAEQLQEEKVSED DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR MAIKFKSFKTKGFGLDACRCMINLMDKDGSG KLGLLEFKILWKKLKKWMDIFRECDQDHSGT LNSYEMRLVIEKAGIKLNNKVMQVLVARYA DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		j	914	ng to first	acid residue	Q=Ghttamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible nucleotide insertion
<del> </del>		<del> </del>	<del></del>	sequence		SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
						EAL
1038	2388	Α	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY
						HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL
		İ		,	1	ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF
		1				GGIETLRVPSELVWLPEIVLENNIDGQFGVAY
			[			DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF   PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG
		1				KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH
1		ľ				GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV
						LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT
						VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
						VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
			1			LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
			,			ELILKKPRSELVFEGQRHRQGTWTAAFCQSL GAAAPEVRCCVDAVNFVAESTRDOEATGEE
1			<b>i</b>			VSDWVRMGNALDNICFWAALVLFSVGSSLIF
						LGAYFNRVPDLPYAPCIQP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL
						QAMENELPVPHTSSSACATSSTSGASSSSGCN
						NSSSGGSGRPTGPQISVYSGIPDRQTVQVIQQ
						ALHRQPSTAAQYLQQMYAAQQQHLMLQTA
						ALQQQHLSSAQLQSLAAVQQASLVSNRQGST SGSNVSAQAPAQSSSINLAASPAAAQLLNRA
						QSVNSAAASGIAQQAVLLGNTSSPALTASQA
						QMYLRAQMLIFTPTATVATVQPELGTGSPAR
						PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL
1010	0000		0.535			PSLALKPTPGGSQPLPTPA
1040	2390	A	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
						HEHRHQSGRCLSTGMAPNLKGRPRKKKPCPQ RRDSFSGVKDSNNNSDGKAVAKVKCEARSA
						LTKPKNNHNCKKVSNEEKPKVAIGEECRADE
						QAFLVALYKYMKERKTPIERIFYLGFKQINLW
1						TMFQAAQKLGGYETITARRQWKHIYDELGG
1						NPGSTSAATCTRRHYERLILPYERFIKGEEDKP
İ						LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP
						KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
						PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA
						PLAPEKDSALVPGASKQPLTSPSALVDSKQES
						KLCCFTESPESEPQEASFPRLPHHTGHRWQTR
L						MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP
	ľ			ľ		GIKARITQRALDYGVQAGMKMIEQMLKEKK LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP
				ļ		NTSLAFVPGVGIKALTNHGTANISTDWGFESP
	ł			ļ	l	LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK
				4		ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE
						NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER
	}		l J	J		SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS
						TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM
		-				VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK
1 1				-	ľ	NSTVETIVSMDFVASTSVGLVILGQRLVCSLS LNRFRLALPESNRSNIEVLRFENILSSILHFGVL
			ĺ		l	PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF
	Ì				ļ	LLISTDLKYETSSKQQPSFHVWEGLNLISRQW
						RGKSAP
1042 ·	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR
						LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  TSLAGNTVQCLNKLKYVIYSAQYPAYGNITT LDMITSTDHVLEQDFWICFTYSVKERQI GLKTRAPATPTFQREVLGPAKQDMQRRCPRI GLMTSLLKPIKRRWRDYKRWKSGGFTGESC HHADTLGDRGGLQGDHSELLQWQKRILRTE GEPSPKYISKNIFPICSYITGFL
1044	2394	A	8718	292	1490	GTVKTSVATPITAGHSCSSGGVLQVKSPATQS GFKFTSKMEDFNMESDSFEDFWKGEDLSNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL VFLLSLLGNSLVMLVILYSRVGRSVTDVYLL NLALADLLFALTLPIWAASKVNGWIFGTFLC KVVSLLKEVNFYSGILLACISVDRYLAIVHA TRTLTQKRYLVKFICLSIWGLSLLLALPVLLFR RTVYSSNVSPACYEDMGNNTANWRMLLRIL PQSFGFIVPLLIMLFCYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM RTQVIQETCERRNHIDRALDATEILGILHSCLN PLIYAFIGQKFRHGLLKILAIHGLISKDSLPKDS RPSFVGSSSGHTSTTL
1045	2395	A	8724	254	3184	FRANLAITVANRRGAQGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFFIGLRGLRGLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVLILNDNLLSSLPNNLFRFVPLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQSDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQQQPPPQLQLQPGEEERRSHHILRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF
1046	2396	A	8736	28	452	SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA
1047	2397	A	8741	673	924	ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL
1048	2398	A	8747	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR

NO: of peptide sequence of the color of the	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Docation   Sequence   1944   1946		NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Docation   Sequence   1944   1946	nucl-	peptide		in	nucleotide	location	
uenice    914   mg to first amino acid residue of peptide sequence   T=Threndine, V=Valiae, V=V=VIPTypophaa, V=V=V=V=V=V=V=V=V=V=V=V=V=V=V=V=V=V=V=	cotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
amino add reidus of sequence peptide seq	seq-	uence	İ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence     Poptide sequence	uence		l	914	ng to first	acid residue	
Poptible sequence   Including insertion   Including insertion   Including insertion   VAVPRIGOPPSAARYMPREVPPRERCQODHX   VAVPRIGOPPSAARYMPREVPPRERCQODHX   VAVPRIGOPPSAARYMPREVPPRERCQODHX   VAVPRIGOPPSSAARYMPREVPPRERCQODHX   VAVPRIGOPPSSCASNING INSPIRITION INTO SEASON   ANSTWOSCASSNING INSPIRITION INTO SEASON   ANSTWOSCASSNING INSPIRITION INTO SEASON   VAKSDPKARS VQSSNISTIENNING ICHWRINYS   GODRIGOGGGGGSTSNING INSPIRITION INTO SEASON   GODRIGOGGGGGGSTSNING INSPIRITION INTO SEASON   VAKSDPKARS VQSSNISTIENNING ICHWRINYS   GODRIGOGGGGGGTSNING INTO SEASON   GODRIGOGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG			İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Incidentide insertion			1	1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
VAVPNGOPPSAARYMPREVPRERQODIK: VLLKRGOPPPSCALLGGGAGPPCTAPGAN PNNAQVTGALLQSESGTAPDSTLOGAAASNY ANSTWGSGASSNNGTSPPRITUDVIVTOGS DMEEWPCLASKUTESSSENTTDNNSASPRGSE KSTLPGSTTSNKGKGSGCQSASSGNECNLGV WKSDPKAKSVQSSNSTTENNNGLGNWRNVS GQDRIGGSGGSSPNPRNSPRAWPALVGGTS RKGALETDNNSSAQVSTVGGTSLEQQSKME NGVVFVSGREQAGLINTDGFKNGANTNSL NGSPPMERNKGMPFGMGLGNTSRSTDAPSQ STGDRKTGSVGSWGAARGPSGTDTVSQGSNS GNNGNNGKEREDSWKGASVQKSTGSKDDS WDNNNSTGGSWPFGDDSNDNTWGGDNK MTSGVSQGEWQPTGSDELKIGEWSGPNQPN STGAWDNQKGBPLLENGCAAACRSSGTDTVSGGSNS GNNGNNGKEREDSWKGASVQKSTGSKDDS WDNNNSTGGSWPFGDSNDNTWGGDNK MTSGVSQGEWQPTGSDELKIGEWSGPNQDN STGAWDNQKGBPLLENGCAAACRSSGTDTVSQGSNS GNNGNNGKEREDSWKGASVQKSTGSKDDS WDNNNSTGGSWPFGDSNDNTWGGDNK MTSGVSQGEWQPTGSDELKIGEWSGPNQDN STGAWDNQKGBPLLENGCAARGPSGTDTVSQGSNS GNNGNNGKEREDSWKGASVQKSTGSKDDS WDNNNSTGGSWPFGDSNDNTWGGDNK MTSGVSQGEWQPTGSDELKIGEWSGPNQDNS STGAWDNQKGBPLLENGCAARGPSGTDVSQGSPNQDNS STGAWDNQKGBPLLENGCAARGPSGTDVSQGSPNQDNS STGAWDNQKGBPLENGCAARGPSGTDVSQGSPNQDNS STGAWDNQKGBPLANGCAARGPSGTDVAGAGNAGAAACAAACAAACAAAAAAAAAAAAAAAAAAAA	1		ł		peptide		/=possible nucleatide deletion, \-possible
VILKRGOPPPPSCMLLGGGAGPPCTAFCAM PNNAQYTGALLQSESGTAPSTILGGAAASNY ANSTWGSCASSINGTSPRPHIWDKVIVDGS DMEWPCLASKUTGSSENTTDNNASAPROSE KSTLPGSTTSNKGKGSQCQSASSGMECNLGV WKSDPKASKVQSSINSTTDNNASAPROSE KSTLPGSTTSNKGKGSQCQSASSGMECNLGV WKSDPKASKVQSSINSTTDNNASAPPALVQEGTS RIGALETDNSNSAQVASTVGQTSEGQSKME MAGVPFVVSGREQAGEINTDGFKAGNITNSL NLSSPPMEMCKOMPFGMGLGNTSSTDAPSQ STODRKTGSVGSWGAARGPSGTDTVSQQSNS GNNCHOKGEREDSVKAGSVQSTGASTSDAPSQ STODRKTGSVGSWGAARGPSGTDTVSQQSNS GNNCHNGKEREDSVKAGSVQSTGASTSDAPSQ WDNNRSTGGSWPGPODSNDNKWGEGNK MTSGVSQGBWKQFTGSDELXIGEWSGPNQPN STGAWDNOKGEPILLENGONAQAPCWGRSS STGSEVEGGTGSNRKAGSSDSHNSGGRSY RFTIPPDCQAVQTILLSRGTDLPRVSNTGWG QTQIKKQTTVMLEEVPRPEKKSDKGTEGWS STGAVDNOKGEPILLENGONAQAPCWGRSS STGSEVEGGTGSNRKAGSSDSHNSGGRSY RFTIPPDCQAVQTILLSRGTDLPRVSNTGWG QTQIKKQTTVMLEEVPRPEKKSDKGTEGWSS SGKNGWGEDQTGYNSNSWGGRGNGWGS SGKNGWGEDQTGYNSNWSGGRGNASWGGG PDEXTFSSWNENSKDGGWGGRGPNGGWS SGKNGWGEDQTGYNSNWSGGGPAPROKGWS SGKNGWGEDQTGYNSNWSGGPAPROKGWG SGKNGWGEDGTTWSNGGNASLASKGGWEGKECKS PAWNETGRQPNSWNRQHQQQPPQQPPPPP PEASGSWGPPPPPONVRYSNSSWSGPQPA TPKDEEPSGWEEPSPGSISRKMDIDDGTSAWG DPNSYNYKNNU, WDNSNSGGPAPPNSKDMGTTDS WEEEDGGVWTTTGSQASSHNSASWGG GKKQMKCSLKGGNDSWMDPLAKQPSNMG LLSGTEDNPSKMDLSVGSJSDKKFDVDKRA MILGDFNDSIMRDRSGGFAPPNSKDMGTTDS GYFEKGGSHGLFONTGAGSSHTVVQFIN SSPSLRAQVPPGISPOVSASMLKQFPNSGLS GLFVVPGLSPQGJAMLSQLFOPGYGSGF PHEDNMYPNALNVGLPDLOTKGPFPGVGSGF SGGMDVGMVGGRSGTSSFRCVGYSMME GLFSVATQESAMHRNGATVAPCKTRGGSPY NOFDTPDDTJGGHTGAGDSWLPAKSPPTNK IGSKSSNASWPPPCGVPWKGGNIDPESDP VYTPGSVLGGTATSPTFVDTDHOLLDRITIGGTA LIKYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBARASTWSDGGSVRPSYWLVLHN LTPQJDGSTLRTICMGHGPLTTFLHATLTQTA LIRYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBAKARQTALHMCVCHNTTLAGTA LIRYSTKQBACHTYNTSSL WGPTYFED PHRMGSFRALLAGCFTPNYSSLSWGFTYFYPCP PHRMGSFRALLAGCFTYPTATTSARA-AGWG LETTQMGSPPLTYTHALLSTTT LPPSSLPPFLQTVDSSSACTLDSFFFFLAPWDS POPGGFFRAGGGPTTPATTSARA-AGWG LETTGMGSPPLATLAGTFTTATHLLLSTTT LPPSSLPPFLQTWQSWSPORGWPTYFSSVVLTH LTCTQMGSPPTYNSSLWGWGPTYFOS PHERMGSFRALLALACHTYDG					sequence		nucleotide insertion
PNNAQYTGALLQSESGTAPDSTLGGAAASNY ANSTWIGSGASSNNGSSPRPHITUNDVIVTOGS DMEEWPCLASKUTESSSENTTDNNSASPROGE KSTLPGSTTSNKOKGSOQCASASSINECNLGV WKSDPKAKSVQSSNSTTENNNGLGWRRIVS GODRIGGGGSGSPNNFNSWPSAWPALVOGGTS RKGALETDNNSSAQVSTWGOTSREQQSKME NAGVAFVVSGRQAQHINTDGPKAGNTINSL LSSPPPMBNKGMPFGMGLGNTSSTDAPSQ STGDRKTGSVOSWGAARGPSGTDTVSQGNS GONGRNGKEREDSWKGASVCKSTGSTDAPSQ STGDRKTGSVOSWGAARGPSGTDTVSQGNS GONGRNGKEREDSWKGASVCKSTGSNSD WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSGEWKQFTGSDELKIGEWSGNPQPN SSTGAWDNQKGHELLENQGNAQAPCWGRSS STGSEVEGGSTGSNNKAGSDSDHNSWGGRN QTQIKGDTVWDIEEVPFEGKSDKGTGGWES AATQTKNSGGWGAPPFPGKSDKGTGWES AATQTKNSGGWGAPPFPGKSDKGGREY AATQTKNSGGWGAPPFPGKSDKGTGWES AATQTKNSGGWGAPPSGNNGWSGWGELS ASTEWAPPKTGAVONGCGROPPOQWS SKIKNGWGEWDQTKNNNWSSASRYSWGG EGGQNEIGTWGNGGASLASKGGWEDCKRS ASTEWAPPKTGAPNSWNNGHQQQQPQCOPPPPQ PEASGSWGGPPPPPFGNNRFNSSWSGCPQPA PKDEEPSGWEESPGDSISKMDDIDGTSAWG DPNSYNYKNNILWDKNSQGGRAPREPPILPTP MTSKASADKSNQDQWGESDQPVTGARPS WEEEEDGGVWNTTGSQGSASSHNSASWGG GKQMKCSLKGNNDSWANPLAKQFSNMG LLSQTDNPSSKMDLSVGSISKMCPTNSKA MLGDPFDIMRKDRSGRPPPNSKMGTTDS GYFEKGGSHGLFONGLSOPPPPPFGNAGARSWGGWG GKQMCSLKGGNNDSWANPLAKQFSNMG LLSQTDNPSSKMDLSVGSISKKCPTVDKRA MALGDPFDIMRKDRSGRPPPNSKMGTTDS GYFEKGGSHGLFONGTARPGCHAPPICAPPT MTSKASADKSNQDQWGESDQPVTGARPBC WEEEEDGGVWNTTGSQGSASSHNSASWQGG GKCMMCSLKGGNNDSWANPLAKQFSNMG LLSQTDNPSSKMDLSVGSISKKCPTVDKRA MALGDPFDIMRKDRSGRPPPNSKMGTTDS GYFEKGGSHGLFONGTARPGCHAPPICAPPT MTSKASADKSNADQWGESDQPVTGARPBC WEEEEDGGVWNTTGSQGSASSHNSASWQGG GKCMMCSLKGGNNDSWANPLAKQFSNMG GKGMCSLKGGRPPPNSCHTNGLSP WEEECDGGVWNTTGSQGSASSHNSASWGG GKGMCSLKGGRPPPNSCHTNGLSP WEECKGGGGLGCAACASSKRCGWGNSSSA MTTPLPFPRSKMDLSOLFQTFGLACQ LLQQQQQQLLDNQKKISQAVRQQCQCQL LLQQQQQQLLDNQKKISQAVRQQCGQCQA RAVVSALQQQQQQQQCQCAACASSSPTVLYHN IGSKSSNASWPPEPQFGYPWKGQNIDPESDP VTTGSVLGGATASTPTLTVQFILA LLYSTKGQAAKATALHACVLGATTLAEF ATDDEVSRT-LAQACPPTPAATPSAPAACWG SKSSNASWPPEPQFQTPWKGANSSSA GGSGGADLAGASIWGPPTNSSSLWQFYTVES PRRMGSPALLFOLLOGGSDSSA GGSGGADLAGASIWGPPTVSSSLWQFYTVES PRRMGSPALLFUTLLLERITHLLLSTFT LPPSSLPFIQTVSSSSACTLDSSFFFLAPWDS QGSSGADLAGASIWGWAYGQGPPTVAATPSAPALAGWG LETTQQCGGAG			-				VAVPNGQPPSAARYMPREVPPRFRCQQDHK -
ANSTWIGGSASSINGTSPIPHIUMDKYIVDEG DIMEMPCIASSINGTISSSENTIDNINGSSPOGE KSTLFGSTTSINGKGSQCQSASSGRECHLQIV WKSDFKAKSVQSSINSTTENNIGLGIWRANV GQDRIGPGSGFSNFMPINSPSAWPALVQEGT RKGALETDNINSSASQVSTVQGTSEQOSKIME NAGVIPVVSGREQAQHINTDGFRNGTINSL MASSPIPMENKGMPFGMGLGNTSTADAPQ STDDRKTGSVGSWGAARGPSGTDTVSGQSNS GDNGNNGKEREBSVKGASVYSTGSTSDAPSQ STDDRKTGSVGSWGAARGPSGTDTVSGQSNS GDNGNNGKEREBSVKGASVYSGTGSKDDS WDNNNRSTGGSWAFGPQDSNDKWGEGNK MTSGVSQGWKQPTGSDELKIGEWSGPNGPN SSTGAWDDQKGHPLENQGNAQAPCWGRSS SSTGSEVEQGSTGSNHEKAGSSDSHNSGRRSY RTHIPDCQAVLQHLSRTDLDFVLSNTGWY QTQIKQDTVWDIEEVPRPEGKSDKGTBCWSS AATQTKNSGGWGDLASSGNGWRSGWGELS AATTEWKDFKNTGGWNDTKNNNSSNWGGG PDEKTFSSWWEIDAFSQNGWASGWGELS AATTEWKDFKNTGGWNDTKNNNSSNWGGGR POEKTFSSWWEIDAFSQNGWGGWGNGWS SGKNGWGEVPQTKNSNWESAASKYVSGW GGGPGGPGGWGGWSGWGGWGGWGGWGGWGGWGGWGGWGGWGGWGGWG			İ	1		*	VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN
DMEEWPCLASKDTESSESNTTDNNSASNPGE KSTLPOSTTSNNGKGSQCQASSGPECNLGV WKSDPKAKSVQSSNSTITENNGLGNWRNYS GODRIGPGSGFSNPNSPNSPAWPALVOEGTS RKGALETDNSNSSAQVSTVGGTSREQOSKEM NAGVNPVVSGEQAQHINTDGPKSQTNISME NAGVNPVVSGEQAQHINTDGPKSQTNISME NAGVNPVVSGEQAQHINTDGPKSQTNISME NAGVNPVVSGEQAQHINTDGPKSQTNISME NAGVNPVVSGEQAQHINTDGPKSQTNISME NAGVNPVNSGEQAQHINTDGPKSQTNISME STODIKTGSVGSWQAARGPSGTDTVSGGSSS GNNGNGKEREDSWKGASVQKSTGSKNSS WDNNNRSTGGSWNFGPQDSNDKKWGEGNY MTSGSVGGWKQFTGSDELKIGEWSGPNOPN SSTGAWDNQKGHPLLENCGNAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSSSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQKKQTVVWDIESPVPREGKSKSHNGGRSS STGAWDNQKGHPLLENCGNAQAPCWGRSS STGAWDNQKGHPLLENCGNAQAPCWGRSS AATQTKNSGGWODAPSGSNQMKSGWGELS AATGTKNSGGWODAPSGSNQMKSGWGELS AATGTKNSGGWODAPSGSNQMKSGWGELS SGKNGWGEVDGTNSNNWESSASKRYSOWG EGQONEIGTWGNGGNNDYKNNSNWSGGGR PDEKTTSSWWENPSKDQGWGGGRQPNOGWS SGKNGWGEVDGTNSNNWESSASKRYSOWG EGGONEIGTWGNGGNNSTNNSNSWSGOPQP PPEASGSWGGTPPPPFGNVKTSNSSWSSOPQP PPEASGSWGGPPPPPFGNVKTSNSSWSSOPQP PPEASGSWGGPPPPPFGNVKTSNSSWSSOPQP PPEASGSWGGPPPPPFGNVKTSNSSWSSOPQP PPEASGSWGPPPPFGNVKTSNSSWSSOPQP PPEASGSWGPPPPFGNVKTSNSSWSSOPQP PPEASGSWGPPPPFGNVKTSNSSWSSOPQP PPEASGSWGPPPFDGNVKTSNSSWSSOPQP QKCMACSLKGGNDSWANDLAGVGGPQCPCOPPPP PEASGSWGPPPFGNSKKMDLSVGGGPAREPHIPT MTSKSADSKSMQDGWGGEGPCFTGARIPS WEEEDGGWWTTTGSGSASSHNSASWGOQG GKKQMKCSLKGGNDSWANDLAGVGGPQCGPCARIPS WEEEDGGWWTTTGSGSASSHNSASWGOQG GKKQMKCSLKGGNDSWANDLAGVGGGG GKKQMKCSLKGGNDSWANDLAGVGOQGQCQ GKKQMKCSLKGGNDSWANDLAGVGOQGQCQ GKKQMKCSLKGGNDSWANDLAGVGGGG GKYQMCGLKGGNDSWANDLAGVGGGG GKYQMCGLKGGAGTGSTRKQGTGGGST GGGMYGMVGGCGCGGMKTSSHPVGTN GFSTATQCAAMMENGALVAPGKTRGGST GGGSGALAGASLWGFTPNSISSPPVGFA RMYSALGQQQQQQQCGMGTMKNHSSSHPVGFA GLPSVATQCAAMMENGALVAPGKTRGGST VYFKSSUGGGSTSGGTLDSFTNVHSTSAK FPDYKSTWSDFIGHNTHLISKNKNHSS RNTTLFRPPGLTNPKPSSWSTAPSWGG GLPSVATQCAAMMENGALVAPGKTRGGST VYFKSSWGGGGRATGGAGTATHMCVCLAGTTLAFA LTYSTTSQAAAASCAGATALMCVCLAGTTLAFA ATDDSVSTELAQAQPPTAATPSAPAAGWG LETCQNQSDPVGPALLEFGGTGLGQWSSSA GGSGAALAGASLWGFPRYSSSWWLINL LTPQDGSTLATLCMGGGSTGLAGTTLAFTPTDTDHQLLADNTTGS PFFLAAWGAAPLLFGGGTGLAGGTGLAGTTLAFTPT LPPSSLPFFLQAWGGAWGP		[		1			PNNAQVTGALLQSESGTAPDSTLGGAAASNY
KSTLPGSTTSNKGKGSQCSASSGNECHLQ   WKSDPKAKSVQSKNSTTENNIGLGNWRNVS   GQDRIGPGSGFSNFMPNSNPSAWPALVQEGET   RKGALETDNNSSASQVSTVQGTSREQGSKME   RAGVNFVVSGREQAQHINTDGFKNGTNTNSL   NLSSPNPMENKGMPFGMGLGNTSSTAPSQ   STODKTGSVGSWGAARGPSGTDTVSGQSNS   GNNGNNGKEREDSWKGASVQKSTGSKNDS    WDNNNSTGGSWNFGPQDSNDNKWGEGNK   MTSGVSQEWKQPTGSDELKIGEWSGPNOPN    SSTGAWDNQKGHPLENQGNAQAPCWGRSS    SSTGSEVEGGSTGSNHKAGSSDSHNSGRSY    RTHPPCQAVLQHLSRTDLDPTVLSNTGWG    QTQIKQDTVWDIEEVPRPEGKSDKGTEOWS    ATTENKEDPKNTGGWDTVKNNSSNWGGGR    PIEKTTSSWNEPSSCQOGGRQPNQGWS    SGKNGWGEEVDQTKNSNWESSASKPVSGWC    DEGOPIEITWENGGANSLASKGWEDCKRS    PANNETGRQPNSWNKQHQQQPPQOPPPO    PEASGSWGGPPPPPFONVRPSNSSWSGPQPA    TPKDEEPSGWEEPSPGSISRKMDIDDGTSAWG    DPNSTNYKNVALWDKNSQGGPAPREPHETT    MTSKASADSKSMQDGWGESDGPTGARHPS    WEEEEDGGVWNTTGSQGSASSHNSASWQG    GKCQMCSLKGGRNDSWMCH_AKQPSNMG    GKQMCSLKGGNDSWMCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GLSQNGCSLGGNDSWMCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWMSCH_AKQPSNMG    GKQMCSLKGGNDSWSGSASSHNSASWQG    GKQMCSLKGGNDSWSGSASSHNSASWGG    GKQMCSLKGGNDSWSGSASSHNSASWGG    GKQMCSLKGGNDSWSGSASSHNSASWGG    GKQMCSLKGGNDSWSGSASSHNSASWGG    GKQMCSLKGGNDSWSGSASSHNSASWGG    GKQMCSLKGGNDSWSGSASSHNSASWGGG    GKQMCSLKGGNDSWSGSASSHNSASWGGGSSASSHNSASWGGGSSASSHNSASWGGGSSASSHNSASWGGGSSASSHNSASWGGGSSASSSSASSHCARGPT    GKCSSNASWFEPQPCVFWKGIONDFESDP    VYTYSSVLGGTATGSFVGVSKAMKGMSSSSACGSSGADLAGASLWGPPTYSSSLWOTPVED    PIRMGSPAPLLFODLLOGGSDSI    GKQMCSLCGGGSDSI    VFWRRQDGSDSSSACCLAGSFFFLAAWTIMLLLS    TAMYGAHAPLLACHTUGAVTLSPTF    LPSSLPFILQAJVTSLACHTUA	1				İ		ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS
WKSDPKAKSVQSSNSTITENNINGLGNWRNYS GORIGPGGGGFNPNSNSPAWPALVQEGTS RKGALETINSNSSAQVSTVGGTSREQQSKME NAGVNVVSGREQAQHINTOEPHSGNTNSL NASSPNPMENKGMPFGMGLGNTSSETDASS STODRKTGSVGSWAARGPGGTDTVSGQSNS GNAGNNGKEREDSVKGASVQKSTGSKNDS WDNNNRSTGGSWAARGPGGTDTVSGQSNS GNAGNNGKEREDSVKGASVQKSTGSKNDS WDNNNRSTGGSWAARGPGGTDTVSGQSNS GNAGNGKEREDSVKGASVQKSTGSKNDS WDNNNRSTGGSWAARGPGSTDTVSGQSNS WDNNNRSTGGSWAARGPGGRYDYGWS STGAWDNQKGHPLLENQGNAQAPCWGRSS STGSEVEGGSTGSNHEAGSSSSNSHNSGRRSS RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQKKQDTVVDDESVPREGKSKNGKGSWS STGKRGWDTVSDEGVPRPEGKSKNGGGS PDEKTPSSWAENPSKDQGWGGGRQPNQWS SGKNGWGEVDTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKS SGKNGWGEVDTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKS SGKNGWGEVDTKNSNWESSASKPVSGWG EGGWRIGTWGNGGNASLASKGGWEDCKS PPEKTPSSWAENPSKDQCWGGGRQPNQWS SGKNGWGEVDTKNSNWESSASKPVSGWG EGGWRIGTWGNGGNASLASKGGWEDCKS PPEKTPSSWAENPSKDQCWGGGRQPNQWG GGKWGGWGGGGNAKGWGGGGRACHAPPTP PRASGSWGGPPPPPGPWRFSNSSWSSGPPOP PRASGSWGGPPPPPGPWFRSNSSWSSGPPOP PRASGSWGPPPPFGPWFRSNSSWSSGPPOP PRASGSWGPPPFPGPWFRSNSSWSSGPPOP PRASGSWGPPPFGPGVFRSNSWSSSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPPGPGNATSNASSWSGPPOP PRASGSWGPPGPGGGGGASASHNASSWGG GKKQMCSLKGGNDSWANPJAACPSNAGG GKCMMCSLKGGNDSWANPJAACPSNAGG GKCMMCSLKGGNDSWANPJAACPGGCQCA LLQQQQQQLLQNQRNGMKHAKSHVGTTNS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLAQVPPQTSGVAGGGNSSWARPPSGPCVFWKGGNJDPESDP VTPGSVLGGTATSPTDTDHQLTRGPPGVAGG SSGGMDYGMVGGKAGTESRFKQWTSMGF GLPSVATQCAMMENGAVAPGKTRGGSPY NGFDIIFGDTLGGTRGPNSTLINKMSNSHSHPVGFK PHLDMWFNALNVGLPDLQTKGPPGVACCA LLQQQQQQQRQFMGMKSNSHSPSTPVGFK PHLDMWFNALNVGLPDLQTKGPPGVACCA LLXYSTSPFDAWPYSASSNSFTNWSTSAS FRYTTLFRPPGLTNKPSSSWSTAPRSVRG GLPSVATQCAMMENGAVAPGKTRGGSSVA LETGQNQSDPVGPALNLFGGSTCLAGWSSSA GGSSGALAGASLWGPFNYSSSLWOPTVED PRAMGSPAPLLFGGTALDFGFTLAGWSSSA GGSSGALAGASLWGPFNYSSSLWOPTVED PRAMGSPAPLLFGGTALDFGFFTLAGWSSA GGSSGALAGAGLWGPTNYSSSLWOPTVED PRAMGSPAPLLFGGTATLAGTTLAGTTLAFF LTDSLYFFLAQAGPTPTAATSPAAAGWS LETGQNQSDPVGPALNLFGGFTLAGTTLAG		1					DMEEWPCIASKDTESSSENTTDNNSASNPGSE
GORLIOPGSGESIAPININSNSAWPALVGGTSREQQSKME RKGALETDNSNSAQVSTVGOTSREQQSKME NAGMIFVVSGREQAQIENTIDGFKINGHTNSI. NLSSPPMPEME KOMPFOMIGLORTSSTDAPSQ STODRXTGSVGSWGAARGPSGTDTVSGQSNS GNNGNINGREREDSWKGASVQKTSVGSQSNS WDNNNRSTGGSWHGPGPGSDILKIGGWSGPNQFN SSTGAWDNQKGPHLENQGNAQAFCWGRSS SSTGSEVEGGSTGSNIKAGSSDSINNSGRRSY RPHPDCQAVLQTILLSRTDLPVLSNTGWG QTQIKQDTVWDIEEVFRPEGKSDKGTGWSG AATQTKNSGGWGAPASGSNMKSGWGELS ASTEWEDPKNTGGWNDYKNNNSNSGGR PDEKTPSSWHGPNGKGRGPROPMS SGKNGWGEDQTTKNNNSSSASKPVSGWG EGQONEIGTWGNGGARGPPNGWS SGKNGWGEDQTTKNNNWSSSASKPVSGWG EGQONEIGTWGNGGARAFCHAFT MYSSASDSKAMDOWGESPOYDFYGAHPS WESEGDGYWNTTGSQGSASSHNSASWGG GKKQMKCSKGGMDDYKNNWSSSASKPVSGWG EGGONEIGTWGNGGASLASKAGNDSWMNHAKAGSNSHM DPNSTYYKNVILWDKNSQGGAPAPEPHLITP MYSSASSSKAMDOWGESPOYTGAHPS WEEEDGGYWNTTGSQGSASSHNSASWGG GKKQMKCSKGKGNDSWMNPLAKQFSNMG LLSQTEDDFPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDSGFPPPSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQFN SSPSILAQVPPQISPQVSASMLKQFPNSGSS GLFNVGPQLSPQJAMLSQLPQIPQFOCACQL LLQQQQQQLQNQKGGKKHSPSHPVGFK PHLDNMYPHALNVGLPDLGYPGOLACQL LLQQQQQQLQNQKISGAWLAGACCL LLQQQQQQQLQNQKGGKKHSPSHPVGFK PHLDNMYPHALNVGLPDLGYPGPQLSCQC LLQQQQQQLQNQKGGKKHSPSHPVGFK PHLDNMYPHALNVGLPDLGYPGCYGCSCG SGGMDYGMVGGKLAGTESRFKQWTSMME GLFSVATQLANNGLAGATASPFVTDIGLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNNHSTAK PPDYSSLYSHGAGYPPQGATATSPAAPAGWQS LETGQNQSTQLATSPFVTDTDIQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNNHSTAK PPDYSSLYSHGAGYPPQATTSPAAPAGWQS LETGQNQSTQLATSPFVTDTDIQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNNHSTAK PPDYSSTYRGAGAPYWGFPPTATSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPAGWQS LETGQNQSTQPAATTSPAAPATTSAAPAGAWQS LETGQNQSTAATSPAATTSPAAPAGWQS LETGQNQSTAATSPAATTSPAAPAGWQS LETGQNQSTAATSPAATTSPAAPAGWQS LETGQNQSTAATSPAATTSPAAPAGWQS LETGQNQSTAATSPAATTSPAAPAGWQS LETGQNQSTAATSPAATTSPAATTSAATTAATTSAATTSAATTAATT		1					KSTLPGSTTSNKGKGSQCQSASSGNECNLGV
RKGALETDNSNSSAQVSTVGOTSREQGSKME NAGWIFVVSGREQQINTORGMINSI. NLSSPIPMENKGMPFGMGINTSI. NLSSPIPMENKGMPFGMGIGHTSRSTDAPSQ STODRXTGSVGSWGAARGPGSTDTVSGQSNS GNNGNMKEREDSWKGASVQKSTGSKMDS WDNNNRSTGGSWHGAGFQDSNDNKWGGSKK MTSGVSQGBWKQFTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS SSTGSVENGGSTGSNIKKAGSSDSINNSGRSY RPTHPDCQAYLQTILSRTDLDPRVLSNTIGWG QTQGKQDTVWDIEVPRPEGKSDKGTEGWES AATOTKNSGGWGDAPSQSNQMKSGWGELS ASTEWEDPSKTDGWNGTGSMDYKSGWGELS ASTEWEDPSKTDGWNGTGSNDFNSSWGGGSG PDEKTTSSWNENPSKDQGWGGGRQPNQGWS SGKAGWGEGEVDQTKNSNWESGWGELS ASTEWEDPSKTDGWNGTSNWGSSWFSSWWG EGGQNEIGTWGNGNASLASKGGWEDCKS PAWNETGRQPNSWNKQHQQQPPQQPPPP PEASGSWGGPPPPFPGNVRSNSSWSGQPPP PEASGSWGGPPPPFPGNVRSNSSWSGQPPPP PEASGSWGGPPPPFPGNVRSNSSWSGQPPP PEASGSWGGPWFPPPFONVRSNSSWSGQPPPP PEASGSWGGPWFPPFONVRSNSSWSGQPPP PEASGSWGGPWFPPFPONVRSNSSWSGQPPPP PEASGSWGGPWFPPFPONVRSNSSWSGQPPP PEASGSWGGPWFPPFPONVRSNSWSGGPPPPP MTSKSASDSKSMQDGWGESDQPVTGARHPS WEEEEDGGWNTTTGSQGSASSHNSASWGQ GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDMPSSKMDLSVGSLSKKPDDKGTSAWG CLSQTEDMPSSKMDLSVGLSCKFRDVKAR NULGDFNDIMKDRSGFPPPNSKDMGTTDS GPYTPEKGGSHGLFGNSTAQSRGLHTPVQFLN SSPSLRAQVPPQFISPQVSASMLKQFPNSCLSP GLINVOPQLSPQQAQQQQQQCGMKHSPSHPVGFK PHLDNNVPNALINVGLPDLQTKGPPGYSGGF SSGGMDYGMVGGKEAGTESRKQWTSMME GLPSVATQEAMMENGALVAGGKTGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSSSSSASWPPPPPGPQFVPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVAPSVWLVLEN LITPQDGSTLRTICMQHGPLLTPHLNKMWKNISS RNTTFLPRPPGLTNFRSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLEN LITPQDDSTLRTICMQHGPLLTPHLNRCMWKNISS RNTTFLPRPPGLTNFRSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLEN LITPQDSTRLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALLILFGTLQGWSSA GGSGADLAQASLWGPPTYSSSLWGPTVED PHEMGSPAPLLAQCPPTPAATTSAPAAGWQS LETGQNQSDPVGPALLILFGTLQGWSSA GGSGAGLAAGASLWGPPTYSSSLWGPTVED PHEMGSPAPLLAQCPPTPAATTSAPAAGWQS LETGQNQSDPVGPALLLFCTLQGWSSACLLDSFFFFLLAPWDS PQDCGFKDHQPLTLQALTVLLARWTIMLLLS TAMYGAHAPLLALCHVDGRVPPRSSACLLT ELTKLLLCASTLLVWGMVAPPPSSSACLLT ELTKLLLCASTLLVWGMVAPPPSPSSACLLT ELTKLLLCASTLLVWGMVAPPPSPSACALLT ELTKLLLCASTLLVWGMVAPPPSPSSACLLT ELTKLLLCASTLLVWGMVAPPPSPSACALLT							WKSDPKAKSVQSSNSTTENNNGLGNWRNVS
NAGWIFVVSGREQAQIENTIDGFKNGHTINSL NLSSPRIMEMIS (MUPPGMIGLOS) STODRXTGSVOSWGAARGPSCTDTVSGQSNS GNNGHNGKEREDSWKGASVQKSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPPQPN SSTGAWDNQKGPHLENGONAQAPCWGRSS SSTGSEVEGQSTGSNIKAGSSDSINSGRRSY RPHPPCQAVIOTILSRTDVSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATOTIKNSGGWGAPASPGNNKSGWGEIS ASTEWKDPKATGGWNDYKNNINSSNWGGR PDEKTPSSWNENPSKDQWGGGRQPNQGWS SGKNGWGEEVDQTKNNINWESASKPVSGWG EGQQNEIGTWGNGGKASLASKGGWEDCKRS PAWNETGRGPNSWNKGHQQOPPQQPPPQ PEASGSWGGPPPPPGNVRPSNSSSSROPA TPKDEPPSGWEEPPPGNSWNKGHQQPPQPPPPQ PEASGSWGGPPPPPGNVRPSNSWSGOPA TPKDEPPSGWEEPPPGNSWNKGHQQPPQPPPPQ PEASGSWGGPPPPPGNVRPSNSWSGOPA TPKDEPPSGWEEPPPGNSKDMGTTDS GPYTEKGSHGLSVGSLSKKFOVDKRA MNLGDPNDIMRKDRSGFRPPNSKDMGTTDS GPYTEKGSHGLFGNSTAGRSGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQPTNSMGS GLTNVGPULSPQOAMSCJOPPOGLACQL LLQQQQQQLDNQKKISQAVRQQCQLA RMVSALQQQQQQQQQQQCQFGRAKEPSHPVGFW PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMFYGNGGKKEPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMFYGNGGKKEPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMFYGNGGKKEPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMFYGNGGKKEPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMFYGNGGKKEPSHPVGFW PHRWSALQQQQQQQQQQCQFGRAKEPSHPVIRGMS IGSSSSASSWPEPPGPQFVPKGQNMIDESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPGPQFVPKGQMGNIDESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPGPQFVPKKGIQNDDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPGPQFVPKGINDDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSSNASWPEPPQFQFVPRWSTAPRSVRG WGTQDSRLASASTWSDQGSVRPSYWLVLHN LITPQDGSTLRTICMQHPPTYSSSLWGYTYED PHRMGSPAPLLFOLLNGRSSYNVHISSAK FPDYSSTWSPDPIGNHPTHLSNKMWKANHISS RNTTTPPPPPPGT-LTPRKTSPYSTYNVHSTSAK FPDYSSTWSPDPIGNHPTHLSNKMWKANHISS RNTTPPPPPFGT-LTPRKTSPYSSTAPRSVRG WGTQDSRLASASTWSDQGSCGLAGVSSASALTLDSFTFLQHWSS PQDCGFKDHQPLTLQALTVELARWTT.MLLIS LTRXYTKQEAACATLHMCVLINTTILLSFTF LPPSSLPFFLQVIDSSSSACTLDSFFPTWSSLALT LETKLLLCASTLLVWGAWPQPPPWRQAA							GQDRIGPGSGFSNFNPNSNPSAWPALVQEGTS
NLSSPPMENKGMPFGMGLGHTSRSTDAPSQ STODRKTGSVOSWGAARGSTODTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS WDNNRSTGGSWNFGQDSDDMKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRS SSTGSVEGGSTGSNIKAGSSDSHNSGRSY RPTHPDCQAYLQTLLSRTDLDPRVLSNTGWG QTQTCCDTVWDTEVPREFERSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS ASTEWEDPKNTGGWNDTSNNSWNGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQOWS SGKOWGEEDDQTKNSNNSSASPVSGWG GGQCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGASLASKPVSGWG EGGCNEIGTWGNGGGGAPREPNLETP MTSKASADSKSMQDGWGESDQPVTGARHPS WEEEEDGGWNTTTGSQGGASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNNG LLSQTEDNPSSKMLDSUGGAPREPNLETP MTSKASADSKSMQDGWGESDQPVTGARHPS WEEEEDGGWNTTTGSQGASSHNSASWGQ GKKQMKCSLKGGNNDSWMNPLAKQFSNNG LLSQTEDNPSSKMLSVGDLACQL LLQQQQQQLLANGKLSGAPPNSCLACQL LLQQQQQQQLLANGKLSGAPPNSCLACQL LLQQQQQQQLLANGKLSGAPPOGCACQL LLQQQQQQQLLANGKLSGAPPYGFK PHLDNNVPNALNVGLPDLOTKGPPGVSGGF SSGGMDYGMVGGKEAGTESRKQWTSMME GLPSVATQEAMMKNGALVAQGKTQCGEQLA RMVSALQQQQQQQCDLACQL LLQQQQQQQLLANGKLSGAPPTYSME GLPSVATQEAMMKNGALVAQGKTGGSSFYN NGDIIPGDTLGGHTGPAGDSWLPAKSPTNK IGSKSSNASWPPEPQGPVYWSGIONDPESDP VYTPGSVLGGTATSPTVDTDHQLLRDNTTGS NSSLATSLPSPGAWPYSASDAVAPGKTRGGSPY NQDDIIPGDTLGGHTGPAGDSWLPAKSPVN LITPQDGSTLATICMQHGPLLTPHLNLTQGTA LIKYSTKQEAAKAQTALHMCGTATUTHLAFF ATDDEVSRLAQAQPPTPAATTSAPAAGWG LETGQNQSDPVGPALNLFGGTGQWSSA GGSGAGLAGASLWGPPTYSSSLWGVPTVED PHEMGSPAPLLPOOLLOGGST. LETGQNQSDPVGPALNLFGTGLQWSSA GGSGAGLAGASLWGPPTYSSSLWGVPTVED PHEMGSPAPLLPOOLLOGGST. LPSTSTPFLQVTDSSACTLDSFFPILLSWTF LPSSLPFPLQTDASSACTLDSFFPILLSWTF LPSSLPFPLQVTDSSACTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPILLSWTF LPSSLPFPLQVTDSSACCTLDSFFPINGAALLT LITPGDGSTLATICHQCHPUTGPPSSSACTLT LETKLLLCASTLLVWGWAPPPYSSACHLT			ł				RKGALETDNSNSSAQVSTVGQTSREQQSKME
STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGREREDSWGAARGVGSTGSKNDS WDNNNRSTGGSWNFGPQDSNDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENGGNAQAPCWGRSS SSTGSVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVIQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQMKSGWGELS ASTEWKDPKNTGGWDDYSNNNSSWWGGGR PDEKTPSSVMEPPSSNQGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPV30WG EGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRGPNSWNKQHQQQPPQOPPPPQ PEASGSWGGPPPPPGNVRSNSSWSGGPPA TPKDEEPSSWEPSSPQSISRKMDIDDGTSAWG DPNSYNYKNVALWDKNSGGGPPREPNLPTP MTSKSASDSSKMDGWGESDOPVTGARHPS WEEEEDGGVWNTTUSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMAPLAKQPSNMG LLSQTEDNPSSKMDLSVGSSAFNSASWGG GKKQMKCSLKGGNNDSWMAPLAKQPSNMG LLSQTEDNPSSKMDLSVGSSAFNSASWGG GKKQMKCSLKGGNNDSWMAPLAKQPSNMG LLSQTEDNPSSKMDLSVGSSAFNSASWGG GKKQMKCSLKGGNNDSWMAPLAKQPSNMG LLSQTEDNPSSKMDLSVGSCFPPNSKDMTIDS GPYPEKGGSHGFGNSTAGSRGLHTPVQPLN SSPSLRAQVPPQFISSQVSSAMLKQPIPQFQLACQL LLQQQQQQLLQNGRKISGAYRQQEQQLA RMVSALQQQQQQQQCQCQQQQQQQQQQQQLACQNGRKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPPGVGSGF SSGGMDYGMVGGKEAGTESFKQWTSMMG GLFSVATQEANMHKNGAIVAPGKTRGGSFY NGFDIIRGDTLGGHTGPAGDSWLPAKSPFTNK IGSKSSNASWPPEPQCVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRNTTGS NSSLNTSLPSFGAWPYSASDNSTTNVHSTSAK FPDYKSTWSPDPIGHTHLSNKMWKNHISS RNTTPLPRFPGLTNYRFSSWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLEN LITQUDGSTLTITACHGHGPLLTFHLITQGTA LIRYSTKGEAKAGATAHMTVLGNTTILAFF ATDDEVSRLAQAQPTPAATTSPAAGWGS LETGQNGSDFVOPFALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSLWGYPTVED PHRMGSFPALFDGLLGGGSDSI VPWKRQDEQLSLQVEIT.YLDSFAVHLLSFTF LPPSSLPFLQVIVDSSSACTLDSFFFFLAPWDS PQDCGFKGDHQPTLTQLATYPLAFWTLLSIT LTAMYGAHAPLLACICHYDGRYPFFRSSAVLLT ELTKLLLCAFSLLVGWAWPQQPPPWRGAA	1		ĺ	[			NAGVNFVVSGREQAQIHNTDGPKNGNTNSL
GNIGNIGEREDSWKGASVQKSTGSKNDS WYDNINRSTGGSWNFEPODSHDNK WGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGEPILLENQGRAQAFUGRESS SSTGSEVEGGSTGSNHKAGSBSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQKQDTVVDIEEVPFFEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGLIS ASTEWKDPKNTGGWNDYKNNHSSNWGGG PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEEVPQTIKNSTWESSASKPVSDWG GGGQNEIGTWGNGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQQPQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP							NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ
WDNNNRSTGGSWNFGPQDSDIDNKWGEGNK MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRS SSTGSEVEGQSTGSNHKAGSSDSINGSRSY RPTHPDCQAVIQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQNMKSGWGELS AATTGKNSGGWGDAPSQNMKSGWGELS ASTEWKDPKNTGGWDDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPYSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS, PAWNETGRQPISWNKQDQOPPQOPPPPP PEASGSWGGPPPPPGNVRSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISAMDDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPYTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQ GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSASSHNSASWGQ GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSASSHNSASWGQ GKKQMKCSLKGGNNDSWMNPLAKQPSNMG GRYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISSPQVSASMLKQPFNSGLSP GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISSPQVSASMLKQPFNSGLSP GLPNVGPQLSPQQIAMLSQLPQPPQFLACQL LLQQQQQQQQRQPGMKHISPSPPVGFR PH.DNMYPNALNVGLEJDTGKPGVASGG SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAJNAPGKTRGGSPY NQFDIPRODITLGGHTTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQGVPWKGGNDFESDP YVTFGSVLGGTATSPIVDTHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNWFSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLINFKPSSPWSSTAFRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLINN LTPQIBGSTLRTICMGHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWG WGTQDSRLASASTWSDGGSVRSYWLVLINN LTPQIBGSTLRTICMGHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWG UCTODSRLASASTWSDGGSUSSSA GGSSGADLAGASLWGPPNYSSLWOYPTVED PHRMGSSPALIFGDLIGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYJDSSVAHLLSPIT LPPSSLPPELDJODSSSAALDSFPSSAVLLL TAMYGAHAPLLALCHYDGRYPFRSSAVLLT TAMYGAHAPLLALCHYDGRYPFRSSAVLLT TAMYGAHAPLLALCHYDGRYPFRSSAVLLT ELTKLLLCAFSLLJWGWAWPQGPPPWKGQA				<b>1</b>			STGDRKTGSVGSWGAARGPSGTDTVSGQSNS
MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDMQKGHPLLENGAQAPCWGRSS SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVINTGWG QTQTKQDTVVDIEEVPREGKSDKGTEGWES AATQTKNISGGWGDASQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWHENPSKDQGWGGGRQPNQGWS SGKNGWGGEVODTKNSNWESSASKPVSGWG EGQQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPQPPPPPPGPPPPP HEASGSWGGPPPPPGNTVRSNSSWSGGPAP PEKAGSWGGPPPPPGNTVRSNSSWSGGPAP PRESEDGGVWNTTGSQGGPAPEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSQGGAPAFEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSQGGAPAFEPNLPTP MTSKSASDSKSMQDGWGESDQPVYGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQ GKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMLSVGSLSDKKFDVDKRA MNLGPFNDINRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSKGLHTPVQPLN SSPSLRAQVPPOFISPVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQPOFOLACOL LLQQQQQQLQRQRQRGMKHSPSHPVGFK PHLDNMYPNALNVGLPDLQTKGPIPGYGSGF SGGMDYGMVGGKEAGTSRFKQWTSMME GLPSVATQEAMMKNGAIVAPGKTRGGSPY NQFDIPGDTLGGHTTPAGTSFFTKAGGSPY NGFDIPGDTLGGHTTPAGTSFFTKAGGSPY NGFDIPGDTLGGHTTPAGNSVLPAKSPPTNK IGSKSSNASWPPEPPQPVPWKGIQNIDPESDP YVTFGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSFGAWPYSASDNSTTNVHSTSAK FPDYKSTWSDPICHNPTHLSNKMWNHISS RNTTPLPRFPGLTNFKPSSFWSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LITPQIBGSTRITCMGHGPLTHH.INJOGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAFF ATDDEVSRFLAQAQPPTPAATPSAAGWGS LETGQNQSPVGPALNLTHTLHLNTLQHTA LIRYSTKQEAAKAQTALHMCVLGNTTILAFF ATDDEVSRFLAQAQPPTPAATPSAAGWGS LETGQNQSPVGPALNLTSPTF LPPSSLPPELDTSLGGGGSDSI VPWKRQDBQLJSLQVETLYLDSSAVHLLSPTF LPPSSLPPELDTSLGGGGSDSI VPWKRQDBQLJSLQVETLYLDSSAVHLLSPTF LPPSSLPPELDTSLGGGGSDSI VPWKRQDBQLJSLQVETLYLDSSAVHLLSPTF LPPSSLPPELDTDSSSSALGGSDSSSA UPPSSLAPGLJGGGGSDSI VPWKRQDBQLJSLQVETLYLDSSAVHLLSPTF LPPSSLPPELDTDDSSSSALGGSDSSSA UPPSSLAPGLJGGGGSDSI VPWKRQDBQLJSLQVETLYLDSSAVHLLSPTF LPPSSLPPELDTDDTSPFPLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRYPFRSSAVLLT ELTKLLLCAFSLLJWGWAWPQGPPPWRQAA	1 .						GNNGNNGKEREDSWKGASVQKSTGSKNDS
SSTGAWDNOKGIPLLENQONAQAPCWGRSS SSTGSEVEGQSTGSINKASSSBHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPREGKSDKGTEGWES AATQTKNSGWGDAPSQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNISSNWGGGR PDEKTPSSWREPPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EOGQNEIGTWGNGGGARQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EOGQNEIGTWGNGGGARQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EOGQNEIGTWGNGGGARGARGARGACH PROEEPSGWEEPSPGSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGAPREENLPTP MTSKSASDSKSMQDGWGESDGPVTGARIPS WEEEEDGGVWNTTOSQGSASSHNSASWGQG GKKQMKCSLKGGNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFPVDKRA MNLGDPNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSGLFGFNSCASGRLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQAMLSQLPQIPOFQLACQL LLQQQQQQQQLLQNGRJSQAVRQQQQQQLA RMVSALQQQQQQQQLQPGFMKHSPSHPVGFR PHDDMWPNALNVGLAGRGGHTPVGSGG SSGGMDYGMVGGKEAGTESRFKQWTSMME GLFSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIEGDTLGGHTGPAGDSWLPAKSPTTNK IGSKSSNASWPPPGPGPVFWKGQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPFRSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRFSYWLVLHN LTQIDGSTLRTICMQHGPLLTFHILNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAFF ATDDEVSRFLAQAQFPTPAATPSAPAAGWQS LETCQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTYASSSLWGVPTVED PHRMGSTRALPGDLLOGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLLDSFTFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLACHZVDGKVPFRPSSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKPPPRSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALALTCHAVDGKVPFRPSSAVLLL TLANGLAGALACHACHACHACHACHACHACHACHACHACHACHACHACHA							WDNNNRSTGGSWNFGPQDSNDNKWGEGNK
SSTGSEVEGÖSTGSNHKAĞSSDSİNISGRRSY RPTHPDCQAYLQTILSRTDLDPRVLSNTGWG QTQIKQDTVWDIEEVPPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS. ASTEWADPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWIEENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGG WEDCKRS. PAWNETGRQPNSWNKQHQQQQPPQPPPQ PEASGSWGGPPPPFGNSRKMODDGTSAWG DPNSYNYKNVILWOKNSQGGPAPEEPILPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGWNTTOSQGASSINSASWGQG GKKQMKCSLKGGRNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MALGDPNDIMRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSCLSP GLFNVGPQLSPQQLAMLSQLPQDFQFQLACQL LLQQQQQQQLQNQKRISQAVRQQQEQQLA RMVSALQQQQQQQQQRQFGMKHSSPSHPVGPK PHLDNMVPRALNVGLPLQTTKGPIPGYGSGF SSGGMDYGMVGGKAGTESRFRQWTSMME GLSSVATQEAMHRIKNGAVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKJPAKSPPTNK IGSKSSNASWPPEPQPGVPWKJPAKSPPTNK IGSKSSNASWPPEPQTGVPWKJPAKSPPTNK IGSKSSNASWPPEPQTGVPWKJPAKSPPTNK IGSKSSNASWPPEPQTGVPWKJPAKSPPTNK IGSKSSNASWPPEPQTGVPWKJPAKSPPTNK IGSKSSNASWPPEPQTGVPKJPAKSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTYQTDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQBAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAGPPTPAATPSAPAAGWQS LETGQNQSPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWQVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSQVETLYLLDSFFFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFPSSAVLLT ELITKLLLCAFSLLVQWQQWPQOPPPWRQAA	1						MTSGVSQGEWKQPTGSDELKJGEWSGPNQPN
RPTHPDCQA'U_OTLLSRTDLDPRU_SNTGWG QTQIKQDTVWDIEEVPPEGKSDKGTEGWES AATQTRKNSGGWGDASQSNQMKSGWGELS AASTEWKDPKNTTGWNDYKNNISNWGGGR PDEKTTSSWNEPNSKDQGWGGGRQPPQGWS SGKNGWGEEVDQTKNSNWESSAKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQQPPPQP PEASGSWGGPPPPPFGNVRFNSNSWSGGPA TPKDEEPSGWEEPSPGSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAREPNL_PTP MTSKSASDSKSMQDGESGDPVTGARHPS WEEEDGGWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMPLAKQFSNNG LLSQTEDNPSSKMDLSVGSLDKKFPVDKRA MNLGDPNDIMRKDRSGFRPPNSCLMSP GFYFEKGGSHG_FGNSTAQSRGLHTPVQLN SSPSLRAQVPPQFSPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQALDLONGKNSQAVRQQEQQLA RMVSALQQQQQQLLONGKNSQAVRQQEQQLA RMVSALQQQQQQLLONGKNSQAVRQQEQQLA RMVSALQQQQQQQQQQCQRGPGMKHSSPHPVGFK PHLDNMVPNALNVGLPDLQTKGPPPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEAMMKNGATVAPGKTRGGSPY NQFDIIPGDTLGGHTQPAGDSWLPAKSPFTNK IGSKSSNASWPPEPQPGYPKKGQNIDPESDP YVTFGSVLGGTATSPTVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTNVHSTSAK PPDVKSTVSPDPJGHNPTBLSNKMWKNHSS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLHN LTQUDGSTLATLSNKGMVKHHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLHN LTQUDGSTLATLSNKGMVKHHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLHN LTQUDGSTLATLGNGGFLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAFF ATDDEVSRFLAQAGPTPAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTPAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTPAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTPAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTPAATSPAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPTPASSELWGVPTVED PHRMGSPAPLLPODLLOGGSDSI  1049 2399 A 8748 200 1387 PWKRQDEQLSLQVETT.YLLDSPATHLAFPTD PPSLPFICLQVDSSSSACTLDSFFFFLAPPUD PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLACHVDGRVPPRPSSAVLLT LETIKLLLCAFSLLVWQQWPQPPPPRRQAA							SSTGAWDNQKGHPLLENQGNAQAPCWGRSS
QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATOTKNSGGWODAPSQSNQMKSGWGELS ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVPQTKNSNWESGASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS. PAWNETGRPYSWKOPQQDPQPPPPQ PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISKMDIDDGTSAWG DPNSYNYKNNULWDKNSQGGPAFERPIN,PTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMPLAKGPSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDPNDIMKNDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPPFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQOLIAMLSQLPQIPOFQLACQL LLQQQQQQLLQNQRKIGAVRQQQEQQLA RMVSALQQQQQQCIJAMLSQLPQIPOFGLACQL LLQQQQQQLLQNQRGGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESFKQWTSMME GLFSVATQEANMHCNGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPPGPGVPWKGIGNIPDESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTAK FPDYXSTWSPDPIGHNPTHLSNKMWKHHISS RNTTPLPRPPPGLTNFKPSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSVNLVLIN LTPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRLAQAGDTPAATTSSALAGWGS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDGI LPSSLPFGLQVIDSSSLACLDSFFPFLAPWDS PHRMGSPAPLLPGDLLGGGSDSL LPSGNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSL LPSSRPFLQLUSSSSACLDSFFPFLAPWDS PQDCGFKDHOPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWKQAA							SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY
ASTEWKDPKNTGGWNDYKNNNSSNWGGGR PPEKTPSSWNENPSSNQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWSSASKPVSGWG EGGORIGTWGMGSMLSASKGGWEDCKRS. PAWNETGRQPNSWNKQHQQQPPQQPPPQ PEASGSWGGPPPPPGNVRENSSWSSGPQA TPKDEEPSGWEEPSQSISKKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTUSQGSASSINSASWGQG GKKQMKCSLKGGNNDSWMPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKA MNLGPFDIMRKDRSGGFPPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQLN SSPSLRAQVPPQFISSQVSASMLKQFPNSGLSP GLFNVGPQLSPQQQQLQNGRKISQAVRQQQEQQLA RMVSALQQQQQQLQNGKISQAVRQQQEQQLA RMVSALQQQQQQLQNGKSSAKKGFNSSHVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMWGGKACTESRFKQWTSMME GLPSVATQEANMHKNGADVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPTINK IGSKSSNASWPPEPQFOVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTIDHQLLRDNTTGS NSSINTSLPSFGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGNPTHLSNKMWKNHISS RNTTPLPRPPGLTNFKPSSPWSSTAPRSVRG WGTQDGSRLASASTWSDGGSVRPSVNLVLHN LTPQGGSTLRTICMQHGPLLTHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGGWSSSA GGSSGADLAGASLWGPPTYSSSLWGYPTVED PHRMGSPAPLLPGDLLOGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVEIT_YLDSPAVHILLSPIT LPSSLBFPFLQVDSSSSACTLDSFTPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPFRPSSAVLTLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG
ASTEWLOPKNITGSWNDYKNNNSSNWGGGR PDEKTFSSWNEIPSDQWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWO EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQQPPQPPPPQ PEASGSWGGPPPPPGNVRFNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNNLWDENSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSINSASWGQG GKKQMKCSLKGGNDSWMMPLAKOPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDPNDIMKNDSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAGSRGLHTPVQPLN SSPSLRAQVPPPGTSSPVSSASMLKOPPNSGLSP GLFNVGPQLSPQQLAMLSQLPQPQFQLACQL LLQQQQQQQLQNQRKJGMLSQLPQPGQLACQL LLQQQQQQQQRQFMKHSSPIPVGFK PHLDNMVPNALNVGLPDLOTKGPIPGYGSGF SSGGMDYGMVGGKAGTESPRKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NOFDIIEGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPPGGVPWKGIGNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWDGGSVRPSVNLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRTLAQAQPPTPAATPSAPAGWGS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLOGGSDSI WYPVKRQDEQLSLQVETLYLDSPAVHILLSPIFL LPSSLBALWGPDNYSSSLWGVPTVED PHRMGSPAPLLPGDLLOGGSDSI WYPVKRQDEQLSLQVETLYLDSPAVHILLSPIFL LPSSLBPFLQVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLTLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
PDEKTPSSWIENPSKDQGWGGGRQPNQGWS SGKNGWGEVDQTKNSNWESSASKPVSGWG EGQQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQPPPP PEASGSWGGPPPPPGNYRPSNSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGAPREPNLPTP MTSKASDSKSMDGWGESDGPVTGARIPS WEEEDGGWNTTIOSQGSASSHNSASWQG GKKQMKCSLKGGNNDSWMPPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDFNDIMRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSCLSP GLFNVGPQLSPQLAMLSQLPQIPQFQLACQL LLQQQQQULQNKISQAVRQQQEQQLA RMVSALQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTIKGPPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLFSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIPGDTL GGTPTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNVBTSAK FPDYKSTWSPDPIGHPFHLSNKMWKNHISS RNTTPLPRPPGLTNPKPSSPWSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAFF ATDDEVSRFLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSFAVIHLLSPTF LPPSSLPFFLQIVDSSSSACTILDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWIT.MLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT				ĺ			AATQTKNSGGWGDAPSQSNQMKSGWGELS
SGKNGWGEFVDQTKNSNWESSASKPVSGWG EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQQPPPPQ PEASGSWGPPPPPGNVRPSNSSWSSGPQPA TPKDEFPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFPPPNSCMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQV3SMLKQFPNSGLSP GLFNVGPQLSFQQLAMLSQLFPQIPQFQLACQL LLQQQQQQQQQQQQQQQQRGFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPCYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLFAKSPPTNK IGSKSSNASWPPEPQFGVWKGIQNIDPESDP VVTPGSVLGGTATSPIVDTDHQLIRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FFDYKSTWSPDPIGHNPTHLSNKMVKNHISS RNTTPLPRPPPGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LITQIDGSTLRTICMQHGFLITFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWGS LETGQNQSDPVGPALNLFGGSTGLGQWSSA GGSSGADLAGASLWGFPNYSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSFAVIHLLSPTF LPPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWITMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSVLLT TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT							ASTEWKDPKNTGGWNDYKNNNSSNWGGGR
EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQPPPPPQ PEASGSWGGPPPPPGNYRFSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVILWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARIPS WEEEEDGGWNTTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKPDVDKRA MNLGDFNDIMRKDRSGFRPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQLAMLSQLPQIPQFGLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIRGDTLGGHTGPAGDSWLPAKSPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNVHSTSAK FPDYKSTWSPDPIGHNFTHLSNKMWKNHISS RNTTPLPRPPFGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAQPPTPAATTSAPAAGWGS LETGQNQSDPVGPAINLFGGSTGLGGWSSAA GGSSGADLAGASLWGPFPYSSLWGVTVED PHRMGSPAPLLPGDLLGGGSDS1  1049 2399 A 8748 200 1387 VPWKRQDEGLSQVETI LYDSPAVIHLLSTFT LPPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWITMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSSVLLT TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT	1						PDEKTPSSWNENPSKDQGWGGGRQPNQGWS
PAWNETGROPNSWNKQHQQQOPPPOPPPO PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA TPKDEEPSGWEEPSPQSISKKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARIPS WEEEEDGGVWNTTGSQGSASSINSASWQG GKKQMKCSLKGGINDSWMPLAKQFSINMG LLSQTEDNPSSKMDLSYGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYTEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPGISPQVSASMLKQPPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNGRISQAWLQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPQYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHINGATVAFKKTGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTINK IGSKSSNASWPPEPGGYPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPFGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGYPTVED PHRMGSPAPLLFGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMILLS TAMYGAHAPLLALCHVDGRVPFFPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
PEASGSWGGPPPPPGNVRFSNSSWSGOPQPA TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWQQ GKKQMKCSLKGGNNDSWMPPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYYEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQQLQNQRKISPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIPGDTLGGHTGAGDSWLPASSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSTTNVHSTSAK FPDYKSTWSPPFIGHNPTHLSNKMVKNHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQDGSTLATCMQHGPLLTHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDYOPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI 1049 2399 A 8748 200 1387 VPWKRQDGGLSLQVETLYLDSPAVHHLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFPRSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							EGGQNEIGTWGNGGNASLASKGGWEDCKRS
TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNYNL WDKNSQGGPARPENLIPT MTSKSASDKSMQDGWGESDGPVTGARIPS WEEEEDGGVWNTTGSQGSASSHNSASWQQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQQQLA RMVSALQQQQQQQQQQQRGMKHSPSHPVGPK PHLDNMYPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIFGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSFGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTTPLPRPPGCLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTLLAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSSPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPIF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP MTSKSASDSKSMQDGWGESDGVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNFSSKNDLSVGSLSDKKFDVDKRA MMLGDFDDMRKDRSGFPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPOPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQEQQLA RMVSALQQQQQQQQRGMKHSPSHPVGFK PHLDNMVPALNVGLPDLQTKGPIFGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIRGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEFQFGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSITWSPDPIGHNFTHLSNKMWKNHISS RNTTPLPRPPPGLTNFKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LITPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGGNQSSPVGAALPLFGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDGDLSLQVETLYLDSSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLYLDSSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTVLGSPAVHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTVLGSPAVHLLSTT LPPSSLPFFLQIVDSSSSACTLDSFFFFLAPWDS PQDCGFKDHQPLTLQALTVLGSPAVHLLSTT LPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVLGSPAVHLLSTT LPSSLPFFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVLGSPAVHLLLST TAMYGAHAPLLALCHVDGRVPFRPSSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTIGSQGSASSHNSASWQOG GKKQMKCSLKGGNNDSWMNPLAKQPSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQEQQLA RMYSALQQQQQQQQQQRQPGMKHSPSHPVGFK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLFSVATACANMHIKNGAIVAPGKTRGGSSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKRSSFWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSFLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVUGRVPFRPSSAVLLT ELITKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
WEEEDGGVWNTTGSQGSASSHNSASWGQG GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MMLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQFGMKHSPSHPVGFK PHLDNNVPNALVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIITGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQFVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPPDPIGHNPTHLSNKMWKNHISS RNTTPLPPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATTSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSFAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPFPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1 :			l		•	
LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQQLQNQRGPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPPRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1					,	GKKQMKCSLKGGNNDSWMNPLAKQFSNMG
GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLQNQRFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFTPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHYDGRVPPRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP GLFNVGPQLSPQQLAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQVVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1						
GLFNVGPQLSPQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQFQVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LIPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPFFLQIVDSSSSACTL.DSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	i				1		
LLQQQQQQLLQNQRKISQAVRQQQEQQLA RMVSALQQQQQQQQRQPGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHLLSPTF LPPSSLPPFLQVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
RMVSALQQQQQQQRQFGMKHSPSHPVGPK PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLCAFSLLVGWQAWPQGPPPWRQAA	1					i	
PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHILSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	]				ļ		
SSGGMDYGMVGGKEAGTESRFKQWTSMME GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMILLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA						l	
GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					ļ		
NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					1		
IGSKSSNASWPPEPQPGVPWKGIQNIDPESDP YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					i	i	
YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LITQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGYPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVHILLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					•	l	
FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPILPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHILSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMILLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					ļ	j	
RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			i				
LTPQDGSTLRTICMQHGPLLTFHLNLTQGTA LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGGLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					ļ		
LIRYSTKQEAAKAQTALHMCVLGNTTILAEF ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA					ł		
ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSLPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
LETGQNQSDPVGPALNLFGGSTGLGQWSSSA GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA							
PHRMGSPAPLLPGDLLGGGSDSI  1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA		. [		1	1		
1049 2399 A 8748 200 1387 VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA				·			
LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1040	2200		0740	200	1202	
PQDCGFKDHQPLTLQALTVELARWTLMLLLS TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA	1049	2399	A	6/48	200	1387	
TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA			.				
ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA						ł	
			.				
						l	
	L		ل	1			PRALDALLY GANNNL VIYLQRYMDPSTYQVL

SEQ ID NO: of nucl- cotide	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq- uence	seq- uence		09/496 914	correspondi ng to first amino acid	corresponding to last amino acid residue of peptide	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
	_			residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				•		SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL MAAGACYAAGGLQVPGNTLPSPPPAAAASP MPLHITPLGLLLLILYCLISGLSSVYTELLMKR QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP
						GLLEGFSGWAALVVLSQALNGLLMSAVMKH GSSTRLFVVSCSLVVNAVLSAVLLRLQLTAA FFLATLLIGLAMRLYYGSR
1050	2400	A	8758	3	1660	WVSSMGFEELLEQVGGFGPFQLRNVALLALP RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS HQDVWLEAHLPREPDGTLSSCLRFAYPQALP
				į		NTTLGEERQSRGELEDEPATVPCSQGWEYDH SEFSSTIATESQWDLVCEQKGLNRAASTFFFA GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV LGLASAASVSYVMFAITRTLTGSALAGFTIIV
						MPLELEWLDVEHRTVAGVLSSTFWTGGVML LALVGYLIRDWRWLLLAVTLPCAPGILSLWW VPESARWLLTQGHVKEAHRYLLHCARLNGR
						PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS
						TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR QTGMGLTALVGRLGGSLAPLAALLDGVWLS LPKLTYGGIALLAAGTALLLPETRQAQLPETI
1051	2401	A	8759	515	1625	QDVERKSAPTSLQEEEMPMKQVQN EIRTPVAVSSAPSGDSEGDEEETTQDEVSSHTS
						EEDGGVVKVEKELENTEQPVGGNEVVEHEV TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV YQHTAAVVSAKSYMCPVCGRALSSPGSLGR
						HLLIHSEDORSNCAVCGARPTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP AGILLVCNNCAAYRKLLEAOTPSVRKWALRR
						QNEPLEVRLQRLERERTAKKSRRDNETPEERE VRRMRDREAKRLQRMQETDEQRARRLQRDR EAMRLKRANETPEKROARLIREREAKRLKRR
1052	2402		07(2	1106	20	LEKMDMMLRAQFGQDPSAMAALAAEMNFF QLPVSGVELDSQLLGKMAFEEQNSSSLH
1032	2402	A	8763	1106	70	RHGHGGRDRRGGGRVARPGGLGRYPGRGAA ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA HGSKHRARAAPDPPPLFDDTSGGYSSQPGGY
						PATGADVAFSVNHLLGDPMANVAMAYGSSI ASHGKDMVHKELHRFVSVSKLKYFFAVDTA YVAKKLGLLVFPYTHQNWEVQYSRDAPLPP
				į		RQDLNAPDLYIPTMAFITYVLLAGMALGIQK RFSPEVLGLCASTALVWVVMEVLALLLGLYL ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL
						LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRQRLQLYLTLGAAAFQPLIIY WLTFHLVR
1053	2403	A	8768	2	712	RPPRVWYPELRELSAAAPRWSHRTAPGIMVF YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE
						VLMDCAHLVKANSIQGCKMNNVNVVYTPW SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER
						NEKKAQIQEMKKREKEEMKKKREMDELRSY SSLMKVENMSSNQDGNDSDEFM
1054	2404	A	8769	344	527	REATTLACRNSCWVFSRCSLGACKPTVCSMP SLSRQGSQTLCLRLAEYCMESVDSQRLLLS

Depuide	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Desire   10   10   10   10   10   10   10   1	NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Seq.   uenice							
Page			ł				
mino seid residue of peptide   residue of peptide   residue of peptide   sequence   peptide	1 .	uaice					
Persidue of peptide sequence	delice	j	J	714			
Peptide			l				
1055						Sequence	
1055			1				
KMI.KCVVVGDGAVGKTCILMSYANDAFPEE   VPPIPPHYAVIVTOVGKQHILGI, VTTAGQ   EDYNOLRPI.SYPNTDYPLICESVVIPASYHNY   QEEWYPELKOCAPPHYVVLIGTODIRDDPK   TLARILYMKEKRITYEHGVYLLAKAIGAQCYL   ECSALTOKGLKA VFDEALITIFHYKKKKRCS   EGHSCSSI    1056   2406   A   8773   261   332   NPRIQLSONSCCAGSCRVWLSEQ   A   8778   3   477   PAGIRIEQARGADMIGKCRGIRTARKIRSH   RRDQKWHDKQYKKAHLGTALKANPFGGAS   HAKGYLEKVOYGKACPNSAIRCKVYOLIK   NOKKITAFYPHOGLIPEENDEVILVAGFGR   KCHANODPOVEREVVXANNSLLAYKGK   KERPRS     1058   2408   A   8808   171   881   POLSQEPSGSMETVVIVAIGVLATIFLASFAAL   VLVCRQRYCRPDILLQRYDSKPIVDLIGAME   TOSSPESILLDDVVITPHHAILENDEWIDA   SGIMSHCIALIKICHTLTEKLVAMTMGSGAK   MKTSASYSDIIVVAKRISPRVDDVVSKSMYPL   DPKILDARTTALLSVSHLVLVTRNACHLTG   GLDWIDGSLSAAGEHEVVIRALALSEPDKG   LPGFGFLQBOGSAI     1059   2409   A   8809   246   757   MELOGASPVLIPHIGPILVWLFTEDHMSGWC   ECPRALSWCFYKVLLLVOTANSVVGYASYL   LVWKDLGGGLGWPLALPIGLYAVQTASYL   VLVFLYFTYSNIDALALLHLLLYGLVVSTALI   WEIDNIKAALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLLEPLA VALUTSAL VLTHAW   UPLIPAKLALLEPLA VALUTSAL VLTHAW   UPLIPAKLALLEPLA VALUTSAL VLTHAW   UP	1055	2405	Δ_	8770		1104	
VYPTYPPHYATTYTYGKQHLLGLYDTAGE	1035	2403	<b>'</b>	0,,0	430	1104	``
EDYNOLRISYMTDVFLICTSVVNPASYHNV	1						
QEEWYPELKDCMPHYPYVLIGTQIDLRDDPK				1			
T.ARILLYMKERPL.TYEHGYKLAKAGIGA,OCY	1		į .				
			ł			ł	
EGHSCCSII	1		ŀ				
1055   2406   A   8773   261   332							
1057   2407   A   8778   3   477	1056	2406	A	8773	261	332	
RRDOKWIDKQYKKAHLGTALKANPFCGAS							
HAKGIVLEKYĞVEAKQPNSAIRKCVRVQLIK   NGKKITAFYPNDGCLNFEDDEVLVAGFOR   KGHAVGDIPGVRFKVVKVANVSLLALYKGK   KGHAVGDIPGVRFKVVKVANVSLLALYKGK   KERPRS     1058   2408   A	100.	2.0.	١	57.75		'''	1
NGKKITAFYPNDGCLNFIEENDEVLVAGFGR			l	ļ			
			l				
NERPRS			ļ				
1058			1				
VLVCRQRYCRRDILQRYDSKPIVDLIGAME   TQSEPSELEIDDVVITNPIEAILENEDWIEDA   SGLMSHCIALKICHTLTEKLVAMTMGSGAK   MKTSASVSDIIVVAKRISPRVDDLVVKSMYPPL   DPKLLDARITALLLSVSHLVLVTRNACH.TG   GLDWIDQSLSAAEEHLEVLREAALASEPDKG   LPGPEGFLQRQSAI   MRLQGALFVLLPHLGPILVWLFTRDHMSGWC   EGPRMLSWCPFYKVLLLVQTATYSVVGYASY   LVWKDLGGGLGWPLALPHGILVAVQTATSVVGYASY   LVWKDLGGGLGWPLALPHGILVAVQTATSVVGYASY   LVWKDLGGGLGWPLALPHGILVAVQTATSVVGYASY   LVWKDLGGGLGWPLALPHGILVAVQTATSVVGYASY   LVWKDLGGGLGWPLALPHGILVAVGTALI   WHPINKLAALLLLPYLAWLTVTSALTYHLWR   DSLCPVHQOPTEKSD   SCKENLLEMWWFQQGLSPLFSALVIWTSAA   FIESYTAVTLHHDPALPHISDTGTVAPEKCLF   GAMLNIAAVLCLATTVAYKQVVALSPEENVI   KLNKAGLVGLIS.CLG.SIVANFQKTLFAA   HVSGAVLTFGMGSLYMFVQTILSYQMQPKH   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLVIWCGYSALSMITCSSVLIN   GKQVFWIRLLLSIDI   GRAPPASVPARESPVSGAQGSSTRGHKRAA   GARAPQLCSSWQRSAPAMSRGLQLLLLSCA   YSLAPATPEVKVACSEDUPCTAPWDPQVP   YTVSWVKLLEGGERMETPQEDHLRGQHYH   QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR   CTTQPPDGQRNLSGKVLRVTGCPAQRKEET   FKKYRAEIVLLLALVIPYLTLIIFTCKFARLQSI   FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT   ELV   CETSTSSAGHAPCRHAAQGPPAFGRRLCSEF   FKKYRAEIVLLLALVIPYLTLIIFTCKFARLQSI   FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT   ELV   CETSTSSAGHAPCRHAAQGPPAGRAEH   CVLSKEMKSVQRSLGLSRHIGPGSPGANTGRALH   CVLSKEMKSVQRSLGLSRHIGPGSPGANTGRALH   CVLSKEMKSVQRSLGLSRHIGPGSPGANTGRALH   CVLSKEMKSVQRSLGLSRHIGPGSPGANTGMIP   LKQVSELSQLYKEAQAELBDYKRRKSLEDV   TAFYHKAEHEKLMQLTNVSRAKAEDALSE   MKSQYSKVLNELTQLKQLVDAQKENSVSITE   MKSQYSKVLNELTQLKQLVDAQKENSVSITE   KKSQYSKVLNELTQLKQLVDAQKENSVSITE   KKSQYSKVLNELTQLKQLVDAQKENSVSITE   LKDTLKSQMTQEAGLADVKRAKEDALSE   KKSQYSKVLNELTQLKQLVDAQKENSVSITE   CKSCHACH   CVLSKENSVSITE   CKSCHACH   CVLSKENSVSITE   CKSCHACH   CVLSKENSVSITE   CKSCHACH   CVLSKENSVSITE   CKSCHACH   CVLSKE	1058	2408	Ā	8808	171	881	
TOSEPSILELIDDVVITNPHEALENEDWIEDA   SGIMSHCIAILKICHTITEKLVAMTMGSGAK   MKTSASVSDIIVVAKRISPRVDDVVKSMYPPI   DPKLLDARTITALLISVSHLVJVTRNACHLITG   GIDWIDQISLSAAEBHLEVLREAALASEPDKG   LPGPEGFI, QGSAI	.000	2400	, · ·	0000	171	001	
SGLMSHCIALLKICHTLTEKL VAMTMGSGAK   MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL   DPKLLDARTTALLISVSHLVLVTRNACHLTG   GLDWIDQSLSAAEEHLEVLREAALASEPDKG   LPGPEGFLQEQSAI   1059   2409   A   8809   246   757   MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC   EGPRMLSWCPFYKVLLLVQTAIYSVVQYASY   LVWKDLGGGLGWFLAPLGLYAVQLTISWT   VLVLFFTVHNPGLALLHLLLLYGLVVSTALI   WHPINKLAALLLLPYLAWLTVTSALTYHLWR   DSLCPVHQPPTEKSD   DSLCPVHQPPTEKSD   VERNEL   V							
MKTSASYSDIVVAKRISPRVDDVVKSMYPPL	1						
DPKLLDARTTALLLSVSHLVLYTRNACHLTG   GLDWIDQSLSAAEEHLEVLREAALASEPDKG   LPGPEGFLQEQSAI			ĺ				
1059   2409   A   8809   246   757   MR.QGAIFVLLPHLGPILVWLFTRDHMSGWC   EGFRMLSWCFFYKVLLLQTATYSVVGYASY   LVWKDLGGGLGWPLAPLGLYAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFTVHNPGLALLHILLLYGLVAVQLTISWT   VLVLFTVHNPGLALLHILLYGLVAVQLTSALTYHLWR   DSLCPVHQPQPGTEKSD   SCKTENLLEMVWFQQGLSFLPSALVIWTSAA   FIFSYTTAVTLHIPAPLPYISDTGTVAPEKCLF   GAMLNIAAVLCIATTYVRYKQVHALSPEENVI   IKLNKAGIVLGIISCLGISIVANFQKTTLFAA   HVSGAVLTFGMGSLYMFQKTILSYQMQPKHH   GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS   GNPGTDLEQKLHWNPEDKGYVLHMITAAE   WSMSPSFFGFFLTYIRDFQKISLRVEANLHGL   TLYDTAPCPINNERTRLLSRDI   TLYDTAPCPINNERTRLLSRDI   GGAPPASVPARESPVSGAQGSSRTRGHKRAA   GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA   YSLAPATPEVKVACSEDVDLPCTAPWDPQVP   YTVSWVKLLEGGERMETPQEDHLRGQHYH   QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR   CTLQPDGQRNLSGKVILRVTGCPAQRKEET   FKKYRAEIVILLAVIPTLTLIPTCKFAALQSI   FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT   ELV   CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE   HQRLHAWPPGPRRPSLWPFKNGKWHSGKRT   AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ   KCSLMCPHRSDGSLSTATIPGRSSPGANTGRALH   CVLSKEMKSVQRSLGLSRIHLQSKRKIHFVL   TR   TR   TR   TR   TR   TR   TR   T	1						
LPGPEGFLQEQSAI							
1059							
EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY	1059	2409	A	8809	246	757	
LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTYHNPGLALLHLLLYGLVVSTALI WHPINKLAALLLLPYLAWITVTSALTYHLWR DSLCPVHQPQPTEKSD  1060 2410 A 8810 304 381 PKLSVYPLQSHHCLSEPFQSLVCCLA 1061 2411 A 8820 1673 848 SCKTENLLEMWWFQQGLSFLFSALVIWTSAA FIFSYTTAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATTVYRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSTSFFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIHFFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYNKRKSLEDV TAEYHKAEHEKLMQLTNYSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1000	2.07	**	000	240	,3,	
VLVI_FTTVHNNGLALLH_LLLVGLVVSTALI   WHPINKLAALLLLPYLAWI_TVTSALTYH_WR DSL_CPVHQPQPTEKSD				1			
WHPINKLAALLLLPYLAWLTVTSALTYHLWR   DSLCPVHQPQPTEKSD	1						
DSLCPVHQPQPTEKSD	1 1						
1060							
1061 2411 A 8820 1673 848 SCKTENILEMWWFQQGLSFLPSALVIWTSAA FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATIYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTILFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI TLYDTAPCPINNERTRLLSRDI GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPFKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1060	2410	A	8810	304	381	
FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATIYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFFFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTIRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYHIKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
GAMLNIAAVLCIATIYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQUYKEAQAELEDYRKRKSLEDV TAEYHKAEHEKLMQLTNVSRAKEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE						* 1.0	
IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTTLSYQMQPKH GKQVFWIRLLLVTWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPFKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSFGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1 ' 1						
HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGK WHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYHIKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1 1						•
GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1						
GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGK VILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	i l						
WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI  1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSL WPPKNGK WHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							WSMSFSFFGFFLTYIRDFOKISLRVEANLHGI.
1062 2412 A 8824 1 763 GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNITISCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE					l		
GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNITISCNSGTYR CTLQDPDGQRNLSGKVILVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSL WPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1062	2412	Ā	8824	1	763	
YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQPPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							7
CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE					}		
FKKYRAEIVLLLALVIFYLTLIIFTCKFARLQSI FFDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE					ļ		
FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV  1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSL WPPKNGK WHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE					1	l	
1063 2413 A 8826 147 627 CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
1063   2413   A   8826   147   627   CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR   1064   2414   A   8835   2982   1869   LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE			Į				
HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1063	2413	Α.	8826	147	627	
AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE			- 1				•
KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE			l	1			
TR  1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
1064 2414 A 8835 2982 1869 LKDTLKSQMTQEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE		1	Į				,
LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE	1064	2414	A	8835	2982	1869	
TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTQLKQLVDAQKENSVSITE							
MKSQYSKVLNELTQLKQLVDAQKENSVSITE			1				
		· ]	J	1	. ]	J	
			ļ				HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	l	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first		Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
3				peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
<b> </b>	•			boquerico	<del></del>	VAKLEKQLLEEKAAMTDAMVPRSSYEKLQS
						SLESEVSVLASKLKESVKEKEKVHSEVVOIRS
						EVSOVKREKENIOTLLKSKEGEVNELLOKFO
						QAQEELAEMKRYSESSSKLEEDKDKKINEMS
1 1	•					KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
1						LQQQVKQLQNQLAECKKQHQEVISVYRMHL
1 .						LYAVQGQMDEDVQKVLKQILTMCKNQSQK
						K
1065	2415	Α	8841	3 .	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
						APLPTGRAQMSPSGRLCLLTTVGLILPTRGQTL
1						KDTTSSSSADATIMDIQVPTRAPDAVYTELQP
						TSPTPTWPADETPQPQTQTQQLEGTDGPLVT DPETHKSTKAAHPTDDTTTLSERPSPSTDVQT
1						DPOTLKPSGFHEDDPFFYDEHTLRKRGLLVA
i			ľ			AVLFITGIILTSGKCRQLSRLCRNHCR
1066	2416	A	8853	3806	2204	FVGEOEGGCEAGAGRGAOTYPGEAGERWFG
						RRRRGRVVSRKKMSLKSERRGIHVDQSDLL
1 1						CKKGCGYYGNPAWQGFCSKCWREEYHKAR
						QKQIQEDWELAERLQREEEEAFASSQSSQGA
						QSLTFSKFEEKKTNEKTRKVTTVKKFFSASSR
}						VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE
						FLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIE
						EQSECAQDFYHNVAERMQTRGKVPPERVEKI
						MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI
1						QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI
						KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI
					i	QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE
					•	KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES
						WSPDACLGVKQMYKNLDLLSQLNERQERIM
						NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI
i						KPPNQPLAAIDSENVENDKLPPPLQPQVYAG
1067	2417	Α	8855	1372	1513	SNMREVGCGWLVPVIPAFWEAEVGGSLEARS
						LRQAWATKQDPISKKK
1068	2418	A	8856	1530	1583	PCRPGMECNSMISVHCNL
1069	2419	A	8857	1530	1583	PCRPGMECNSMISVHCNL
1070	2420	Α	8866	293	1675	PYPQGGYPQGPYPQGGYPQGP
						YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ
			1			EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF
	İ					LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
			-			WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG
1		.				ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM
		ļ	1	ſ	1	VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA
		1		ļ		VDTOLLLGNKOLSLSPEEYVFAALNLYTDIINI
1 1	ľ	i	ľ	- 1	i	FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP
			ł	İ	ļ	WHGSASCTSPLSCPQAQPREKDASLQPSCMY
			j			TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC
		ļ			}	HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ
	1	ĺ	ĺ		·	EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS
						GDMRSGGLIPVLSPE
1071	2421	Α	8868	2	358	ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH
1	j	J	j	ļ	ļ	DDKMGSNTFFKRNDCRYVMISCKADMAYDN
		Ì				VRHPFMI*SI\KLIMEETYLNIIKAVYDRPTASII
1075	2422	$\rightarrow$	0070		750	LNGEKLKVFPVRSGT*QGCSVWP
1072	2422	A	8870	33	658	MESVLSKYEDQITIFTDYLEEYPDTDELVWIL
1	]		İ			GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH
						LGRDWSWEKQKEQPKEYQRILQCFLDRKDC
L						DOWN TO THE TOWN THE TOWN OF T

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						CYSIHQMAQMGVGEGKSIGEWVLGPNTVAQ GV*KNLA\LFDEW\NSLGLVYVSM\DNPSGSIA RFPKKLCRVLPL\SADTAGLTGP
1073	2423	A	8879	146	412	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL *QVCITA*IKESVIISGG*SSSPVCHTTFQPANL RTSRYLPT*SIKSLGPDEPQEG
1074	2424	A	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW KEISFGDYICHTFQGDCWADRSPLHEAAAHG RLLALKTLIAQGVNVNLWTL/DRVSSLHEACL *GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	A	88%	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRS SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK PWPSLLDKEREESLRQKRLSERERIGELGAPE VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS TSEEEKKKKSSRSKERSKKRKKKSSKRKHK KYSEDSDSDSDSETDSSDEDNKRRAKKAKK KYSEDSDSDSDSETDSSDEDNKRRAKKAKK EKKKKHRSKKYKKKRSKKSRKESSDSSSKES QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGEGAAMAEYVKAGKRI PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR MEAVRTAKREPESTVLMRREPLHPFNPRRET KERE
1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E *APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE VSYIVSSREVKAESSGKSHRGCPSPSPSEVR VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	Α	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW QYPALHRAGTEWQLSALHRAPRSTQPDKAC RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT \YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT
1079	2429	A	8912	121	376	NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY
1080	2430	A .	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGYFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC
1081	2431	A	8922	56	420	EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1082	2432	Α	8923	355	1079	LSIAIPAMVNNTAPPSQPNASTERPST PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

Decide   Seq	SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
Sequence			nou				
1914							
minto acid residue of peptide residue of peptide residue of peptide sequence   T-Threonine, V-Valine, W-Tyroponian, V-Sup codon, P-possible nucleotide deletion, V-possible nucleotide deletion, V-possible nucleotide deletion, V-possible nucleotide insertion   GPPSKEDQYDVKAPAAFNIRNIGK/TLVART QOTIOLASQI, GGLEVEVIALDI, ONDEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, ONDEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, ONDEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, ONDEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, ONDEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVEVIALDI, GGLEVARY REMAINED   GOTIOLASQI, GGLEVARY REMAINED   GGLEVARY GGLEVAR	1 7						
Peptide   Sequence   Peptide   Sequence   Peptide   Pe		Ì	l				
Peptide	1	1					
					peptide	1	
GPFSKKDQYDYKAPAMPNRNTGK/TLVART		f	[	i			
QGTQLASDGLKGLLFEVSLADLQNDEVAFRK   FKLITEDVQDENCL.ThYGMDLTGCKIGSMV							GPFSKKDQYDVKAPAMFNIRNTGK/TLVART
FKLITEDVQDKNCLTNFYGMDLTCDMICTOM	1	ŀ					OGTOIASDGLKGLLFEVSLADLONDEVAFRK
HNNQLKTYSYA-HQGSRQIQKKMMEINTEN	1	ļ	1				FKLITEDVQDKNCLTNFYGMDLTCDKICSMV
HNNQLKTYSYA-HQGSRQIQKKMMEINTEN	1	ł	1				EKWSTMIEAHVDVKTTDGYFFHLFCVGFTKK
			İ				
1083	]	}				ł	QTNDLKEVVNKLIPDNIGKDTEKV/CPIYPLH
WGAVQGSRAMSDLILLUTLILLUTLLILLUTLLILLUTLLILLUTLLILLUTLILLUTLLILLUTLILUTLILLUTLILutlilutlilutlilutlilutlilutlilutlilutli			1	]			
AGYSGQLAGVAVSAGSPPIRYKFHVEYYGET	1083	2433	A	8948	28	385	LTWPQPHIPSCPAMSEETLQSKLAAAKKKLP
1084   2434   A   8950   156   318				1			WGAVQGSRAMSDLLLLLLDLTLLLLLMLLGF
1084		Ì		ļ			AGYSGQLAGVAVSAGSPPI/RYKFHVEPYGET
1085   2435   A   8956   16   413   HMGQLGYFIQCWWECKILISHWKTI*QSPAK	1		1				GWLLT/ESCSISPKLCSIAVH*DNPAWF
1085	1084	2434	A.	8950	156	318	HYTPINTDTIENSENNKCW*GY*E\VGLIHHW
**TIYTSYDTAPIS/GIVPKRMSSKCHOETCAR   MFILAPFTATIKGKQLTCPLVERIDYMWYS   HKYYIKVKRRL*VTITHATIWNINIMEBILW   YSHKYY   HYTHATIWNINIMEBILW   YSHKYY   HKYYIKVKRRL*SQLLRRLRHESHL   NPGARGCSEARLHRCTPAWTT   NPGARGCSEARLHRCTPAWTT   NPGARGCSEARLHRCTPAWTT   1087   2437   A   8985   58   330							WGGKRVQPFWKRVWQKRTLNLRV
MIPILAPFTATIKGKQLTCPLVEERIDYMMYYS	1085	2435	A	8956	16	413	HMGQLGYFIQCWWECKRLISF\WKTI*QSPAK
HKYYIKVRRNL*VTITH\text{TWNLNILMFEILW}							*TIYTSYDTAIPIS/GI/YPKRMSSKCHQETCAR
1086			ł				MFILAPFTATIKGKQLTCPLVEERIDY\MWYS
1086			l				
1087   2437   A   8985   58   330	L						
1087	1086	2436	A	8962	868	1026	
*ERSVCAFHVCIQTYVCLQVYACMCVYYICM FYYSVYGCGLCTCVCMDVYICVCQFEL							L
FYYSYYGGLCTCVCMDYYICVCVQEFL	1087	2437	A	8985	58	330	
1088	1 .	j ,	,				*ERSVCAFHVCIQTYVCLQVYACMCVYYICM
KKFLDMMSNA*STKKHDKLD/LIKFKT/LCSA   KYTYKRIKHHPTDLEKMLRNHLSDKD*YS/GV   YKDLSKLNRKTE/S*/VKKWVKDLSRYFIKE   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHKKIFSTS   VISMENKHV   VISMENKH   VISMENKHV   VIS		2100		2000			
KYTVKRIKIHPTDLEKMLRNHLSDKD*YS/GV	1088	2438	A	8989	394	404	
1089   2439   A   8991   60   329   MALTPESPSFFGLAATGSSVPEPPGFPNATL   NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV   GVEDNAYTLEVNSRYMRAVGIM*IHL   NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV   GVEDNAYTLEVNSRYMRAVGIM*IHL   SNITTLT*MKKYDNITEV*GCGGGJT/LIYC   WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL   GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR   LLHALFILVQYWKQSKGHYI   LULAALFILVQYWKQSKGHYI   LVSSRFKISKVIVVGGLSVGKTCLINR*GGAG   AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS   FDKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVLGIPF   STKNYKAPIGADFEMERFEVLGIPF   STKNYKAPIGADFEMERFEVLGIPF   SPKNYKAPIGADFEMERFEVPICTOSPEPRQEVPMCTOSPEPRQE			1			*	
VISMENKHKKIFSTS	1						
1089   2439   A   8991   60   329   MALTPESPSSFPGLAATGSSVPEPPGGPNATL   NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV   GVEDNAYTLEVNSRYMRAVGIM*IHL	i i						
NSSWDSPTEPSSLEDLEATGTIGTLLSDMGVV GVEDNAYTLEVNSRYMRAVGIM*IHL  1090 2440 A 8996 2 351 SNITITLT*MKXYDNTFCW*GCGQIG/T/LIYC WQESKFIQAFWSKIQQYLA*ISHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQIVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHLIFEDDWHQTTCCHHPHHPV*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHQQPLGHGLVARVG  1093 2443 A 9002 3 2745 ALGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEELPI_VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV*LWTINK*YP RVQTTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPMCTGSEPRQEVPM GGKFPCIHI/PEPRQEVPMCTGSEPRQEVPM MCTGPEARQEVPMCTDSEPRQEVPMCTGFEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTGFESRQEVPM	1089	2439	Δ	8001	60	329	
GVEDNAYTLEVNSRYMRAVGIM*IHL  1090 2440 A 8996 2 351 SNITTIT.*MKKYDNTFCW*GCGQIGT/ILIYC WQESKFIQAFWSKIQQYLA*ISHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLFT LVSSRFKISKVIVVGDLSVGKTCLNR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPV*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAHHHQHGQGFLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHCTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLQGETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFKV*LWTNK*YP RNGIISFTRT/INLHGFSKNPKV*LWTNK*YP RNGIISFTRT/INLHGFSKNPKV*LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPMCTOSPEPRQEVPMCTOSPEPRQEVPM CTGPPERQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMYTGSEPRQEVPMYTGSERVEPVMY	1.00	2437	11	6,,,,	00	327	
1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCW*GCGQIG/T/LIYC WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLLAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQWCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAGSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELABHHRSK VPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISLCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESRFYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCTDSEPRQEVP	}						
WQESKFIQAFWSKIQQYLA*ISIHILFDPAFLFL GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPGEHWNSLGGKSSCPLPLPT LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG APLGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIRHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFTYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQQPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLQGETLCSLQEADYEVASYGLQ HNCLGILGGDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPMCTDSEPRQEVPPMCTDSEPRQEVPMCTDSEP	1090	2440	Α	8996	2	351	
GGYPGGTQSVFLTGVLVSSVFYNMKMLHTR LLIAALFIIVQYWKQSKDHYI  1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILFEDDWHQTTCCHHPHHPF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAGSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLVLYQGEKKLEELPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCHI/PEPRQEVPTCSDPEPRQEVPMCTGSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMCTDSEP			•				
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIBF  1092 2442 A 8999 548 811 SSFIKHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPERPGEVPTCTG PESRREVPMCSDPEPRQEVPTCTDSEP RQEVPMYTGSEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMTTGPESRQEVPMY							
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWNSLGGKSCPLPLPT LVSSRFKISK VIVVGDLSVGKTCLINR*GGAG AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIBF  1092 2442 A 8999 548 811 SSFIKHILIFEDDWHQTTCCHHPHHPVF*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALGLQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPERPGEVPTCTG PESRREVPMCSDPEPRQEVPTCTDSEP RQEVPMYTGSEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMTTGPESRQEVPMY							LLIAALFIIVQYWKQSKDHYI
AELGRVGPSLARWAGSRSQHLVPSQVCKDS FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPV*RCQ FHIFYVSVQNSISPSLSVSSSHDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFPSGLA VFTRFALKTLQGETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1091	2441	Α	8997	97	456	
FDKNYKAPIGADFEMERFEVLGIPF  1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPY*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTTVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEPRQEVPMY							LVSSRFKISKVIVVGDLSVGKTCLINR*GGAG
1092 2442 A 8999 548 811 SSFIKRHILIFEDDWHQTTCCHHPHHPY*RCQ FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1						AELGRVGPSLARWAGSRSQHLVPSQ\VCKDS
FHIFYVSVQNSISPSLSVSSSHPDRPDHEVHQH RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY					·		FDKNYKAPIGADFEMERFEVLGIPF
RAAHHHQHGQGPLGHGLVARVG  1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMCTDSEP	1092	2442	A	8999	548	811	
1093 2443 A 9002 3 2745 ALLGLQQPAQ\$LILSRSSVMGVRGLQGFVGS TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
TCPHICTVVNFKELAEHHRSKYPGCTPTIVVD AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
AMCCLRYWYTPESWICGGQWREYFSALRDF VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1093	2443	A	9002	3	2745	
VKTFTAAGIKLIFFFDGMVEQDKRDEWVKRR LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTINK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
LKNNREISRIFHYIKSHKEQPGRNMFFIPSGLA VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTINK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	į				:		· · · · · · · · · · · · · · · · · · ·
VFTRFALKTLGQETLCSLQEADYEVASYGLQ HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
HNCLGILGEDTDYLIYDTCPYFSISELCLESLD TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPLVTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTINK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1 1					l	
TVMLCREKLCESLGLCVADLPLLACLLGNDII PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
PEGMFESFRYKCLSSYTSVKENFDKKGNIILA VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	ļ l						
VSDHISKVLYLYQGEKKLEEILPL/VTKQSSFL *RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY	1				l		
*RNGIISFTRT/INLHGFSKNPKV**LWTNK*YP RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY					l	ļ	
RVQTPNPGKKFPCVQMLNPGKKFPCVQALNP GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY					l		
GEKFPCIHI/PEPRQEVPTCSDPEPRQEVPTCTG PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY							
PESRREVPMCSDPEPRQEVPMCTGPEPRQEVP MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		-				l	
MCTGPEARQEVPMCTDSEPRQEVPMCTDSEP RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY		İ	ľ		l	Ì	PESBENDMCSDPEPROFUPMCTGPEPPOEVD
RQEVPMYTGSEPRQEVPMYTGPESRQEVPMY					l		MCTGPEAROEVPMCTDSEPROEVPMCTDSEP
	1 1	ł	ļ	. 1	ł		TGPESRQEVLIRTDPESRQEIMCTGHESKQEV

SEQ ID NO: of nucl- eotide	NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue	Q=Glutarnine, R=Arginine, S=Serine,
1				residue of	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon.
i	1	1		peptide	Supulation	/=possible nucleotide deletion, \=possible
		l		sequence		nucleotide insertion
		<del> </del>	$\overline{}$			PICTOPISKQEDSMCTHAEINQKLPVATDFEFK
}				İ		LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS
1		1	1		[	DTEILKVARTHHVQAESYLVYNIMSSGEIECS
		}			Ì	NTLEDELDQALPSQAFIYRPIRQRVYSLLLED
		1	ì		ļ	CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL
			į .			QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
1	1	i				CFNLSSSREELQAVESPFQALCCLLIYLFVQV DTLCLEDLHAFIAQALCLQGKSTSQLVNLQP
1		Ĭ				DYINPRAVQLGSLLVRGLTTLVLVNSACGFP
						WKTSDFMPWNVFDGKLFHQKYLQSEKGYA
						VEVL/CRTK*ISAHQIPQPEGSRLQGLHEGEQT
i						HHWPSPLGLTPRREVGKTGLQLPQDGLWV
1094	2444	Α	9021	97	834	AREACRAKTDFPGRRFRLWPSCCCRVIVGAE
		ĺ				T*H\MAEPVSPLKHFVLAKKAITAIFDQLLEFV
	İ					TEGSHFVEATYKNPELDRIATEDDLVEMQGY
1.	-	ļ				KDKLSIIGEVLSRRHMKVAFFGRTSSGKSSVI
1						NAMLWDKVLPSGIGHITNCFLSVEGTDGDKA
į						YLMTEGSDEKKSVKTVNQLAHALHMDKDLK AGCLVRVFWPKAKCALLRDDLVLVDGPGTD
1						VTTELDSWIDKFCTKSSTREITNSGSDT
1095	2445	Α	9022	1	537	LVLNSRVEDFVPPEGAGRTLPFALRPLAACW
				_		LLHRRARRSSALCPRPRSWGVSGGEGAGARE
						P*ITSSSCCLSAA/SHLSIQSPNMAGARRIRPQ
1	1					LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF
						DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV
1096	2446	_	0020		206	FAYGQT\GAGKTYTMGTGFD
1096	2440	Α	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL
						GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ EDCLSPGGGNCRES*SCPYTPAWITERDPV
1097	2447	A	9032	716	357	ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI
		, ,				LAHCNFRHAGFPPLSCLSLPNRWEYRRPPARP
1						GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG
						LPKCWDYRREPAASIIFQTTFFINSK
1098	2448	Α	9038	230	652	KVVVMSCEDINISGSFYRNKLKYLAFLCKRTS
] .						TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ
1			1			G*AALDHLKVFDRIPLPYDKKKQMAVSATLE
						VVRPKP*RKFAYLGHWAQKVDWKYQAMTA
1099	2449	Α	9043	185	372	TMGEKRKVYYQKICYQKK  IFYSHQQCMRV/WQGCGDIETLIHCW*E*KII
1	~		1043	105	J/2	HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP
						RKIKTCPONSCTSMLINAIHNDOKWKKINI
1100	2450	Α	9045	763	584	RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL
	<u> </u>					SLTSSWDYRRPPPHPANFLYFK*RRGF
1101	2451	Α	9050	275	2	LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL
	[ ]		l	l		FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS
1305	1					DSPASAS*VAGITDMCRYTQLILFHAS
1102	2452	A	9053	449	1224	KTSMFWKFDLHSSSHIDTLLEREDVTLKELM
]						DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF
1						IN EEPPQDMDEKIRYKYPNISCELLTSDVSQM
1	1 1			1	ł	NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG
1						TSAIMDLLLRLLTCIEPPOPRODVLN/WFKVO
1						RNL*HST*NVMDISKYVNLHWGLNKSHSLL*
				j		LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS
L	1	- 1	1			SLV
1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
,		ļ				WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA
{				•		AGMQI/H/CW\WCVNVGKFWEMS*YYLLKLSI
L <u>.</u>						ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  APFVLAVNC
				75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V KTDCGCGANSKGVVVVMKV\KTAQQKQTTS YMQIGTTKNSRAT
1105	2455	A	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR AWPCCPGWSAAWLTIVILAHYRRPGLERSCC LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL VLNS*TQGI
1106	2456	A	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT* AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	A	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA SAFPPAERSRGHRRASL*RARWSAAVPRRSA GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD*GG QRPPPPSGDSLSPPGCCRY
1109	2459	Α	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL GAVAHSCNPSTLVGRGGRITRGOELR
1110	2460	A	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP AAAGDPASLDFAQCLGYYGYSKFGNNNNYM NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK RAAPDSGKKPKTPKK
1113	2463	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH VGQAGHEPLTSGDPPASASQSAGITGVSHQA WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS LLKCWDY
1114	2464	A	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII YLYVQTVPQRV
1115	2465	A	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE WRLQHKGRGRGDLHLPDHHLSVPSSADHPA QQPSQFNGRNLYFLPLFR
1116	2466	A	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA QHVHVPPWTDVLAGQDRRAPTAGDGAPWP APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA GAGGP*GSPAGRACGAAGCRPRPPRPAASSA *NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTITTELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG*GILTKLSAFLTIPRLQPHLIAALSPSS
1118	2468	Α	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2469	A	9155	124	207	ACPRLARRRRVSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLQFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGSVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLP\WQRRGLLRAQG\LRG\WKARRA\TTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEDEDDEDGGEEAPAPG GAGKSEGSTPADGLYGNSRHSCFPVV PPRACRPCPRACPCPPT*KCSOPVSWPC
1121	2471	Ä	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
						V/CSHITDSLKFIGKGWVGMVTHACNPGTLG G*GGWIA*VREFETSLGNM
1122	2472	С	9170	442	236	MNRRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHILKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1124	2474	A	9173	3	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL WEDQGGLLGPFSFLMLMLLLETRNPVNACLL TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	Α	9179	704	188	ESSSGLLFQCFQGIHVQKLTLQARPTLFSWWL CSKPPKETGELENAESGGDGGRRGGKQDNV AWWRRMQKG\DFPWDDEDFPQSGPFGGQA LPMGFFYLYFRDPGREITWKHFVQYYLARGL VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE YCRHKFISCKNVVFYFFQ
1126	2476	A	9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	A	9185	1	321	LTGQLGSILLRVFSKSRAGLGARKLKAYRTM EYMAESTDRSPGHILCCECGVPISPNPAQY\CV ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLQTMDEQSQGMQGPPVPQFQPQKAL RPDMGYNTLANFRIEKKIGRGQVFSEVYRAAC LVLDGVPVALKKVQIFDLMDAKARADCIKEID LLKQLNHPNVIKYYASFIEDNELNIVLELADA GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS ALEHMHSRRVMHRDIKPANVFITATGVVKLG DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD NG
1129	2479	A	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT RATT\KIRVVATTTRARIEDMRHSATALTRPD ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A .	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE PLAFLRALELLFAIFAFATCGGYSGGLRLSVD CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIQLTTK TQPVEATDDAFWDQFWADTATSVQDVFALV PAAEIRAVREESPSNLATLCYKAVEKLVQGA ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD WRGFFWSTVPGAGGGGEEDDEHARPLAE SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH SLDSCEYIWEAGVGFAHSPQPNYIHDMNRME LLKLLLTCFSEAMYLPPAPESWQHRTHWFSS FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY NHLY
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS NVYFIV
1134	2484	A	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK RADGAAPAGEGEGVTLQGNITLLKGVAVIVV AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC GVFSIVGALCYAELGTTISKSGGDYAYMLDV YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC YSVKAATRVQDAFAAAKLLALIILLGFVQI GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF
1 1						GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS
					}	RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT
						CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII
1						GMIWLRHRKPELERPIKVNLALPVFFILACLF   LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV
						WWKNKPKWAPPGHLSPRPSCVRSSCMVVPO
1135	2485	A	9216	40	410	RDRLPPAYFCRPVVCVVTALDVG\SPESOEM
		•	,2.0			DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV
						MQETLRNLASIGEKWKDQNIEDQYKNPRNNL
]						RSLLGERVDENTEENHCGETSSQIPDDTLNK
1136	2486	Α	9223	3	983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG
		Ì				DLVFAKMKGYPHWPARIDDIADGAVKPPPN
						KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK
1 1						PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD
						SEAPEANPADGSDADEDDEG\RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK
						MSVSKRARKASSDLDQASVSPSEEENSESSSE
						SEKTSDODFTPEKKAAVRAPRRGPLGGRKKK
1						APSASDSDSKADSDGAKPEPVAMARSASSSSS
.						SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK
						PKPERPPSSSSSD
1137	2487	Α	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASQV
		ľ				GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL
1138	2488	A	9231	1664	2	RGSREPPAWA
1136	2400	A	9231	1004	2	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL EGIVWHETEEGVLVVNVTWRNKTYVGTLLD
						CTKHDWAPPRFCESPTSDLEMRGGRGRGKR
						ARSAAAAPGSEASFTESRGLQNKNRGGANGK
1						GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR
						KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP
						TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK
						YKHINGLRYHQAHAHLDPENKLEFEPDSEDK
		1				ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA
1						SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK NSGKKKGLNNELNNLPVISNMTAALDSCSAA
				'		DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK
1						ANNCKTDKN\PSKLKSARPIAPAPAPTPPQLIA
]						IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL
		1				KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR
						KLKDKEGKETGSPKMDAKLGKLEDSKGASK
						DLPGHFLKDHLNKNEGLANGLSESQESRMAS
1100	2405				448	IKAEADKVYTFTDNAPSPSIGS
1139	2489	A	9234	207	443	TRRGQPWRRRAAAAGILPGREAAACLPSC/AS
		1				VTAAVSGLLVGYELGIISGALLQIKTLLALSC
1140	2490	<del>_                                    </del>	9238	248	328	HEQEMGVSSLVIGALL MAQGNNYGQTSNGVADESPNMLVYRKV
1140	2490	A	9236	240	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP
1171	471	^	7272	-	555	TLRGHGGASGRNVTTGSLGEPOWLRVATGG
				ĺ		RPGTSPALFSGRGAATGGROGGRFDTKCLAA
		l				ATWGRLPGPEETLPGQDSWNGVPSRAGLGM\
	*	İ				WPWAAALVVHCYSKSPSNKDAALLEAARAQ
						VMQEVSRNRCALLHSAAVQEYGYGN
1142	2492	A	9245	157	466	HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD
						FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC
						CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF
	I	1	1			OYOURSONA
11/2	2402		0247	264	116	GICKEYSRQ
1143	2493	A	9247	264	115	GICKEYSRQ GLPQQTSTIQPPGTPDGARDFTSTIQPPGAPDG ARDSTSIIRMGPEIPPP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE AMEESDRPCEISEIDDNPKISENPRRSPTHEKN TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM ERRR
1145	2495	A	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL WDTAGQERFISIT
1146	2496	A	9277	592	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA WLIFAFLVETGF
1147	2497	A	9279	1255		FRRGRRGEEKEEEEEEEGWVNGMENSHPP HHHHQQPPPQPGPSGERRNHHWRSYKLMIDP ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKELELSVPKFKIDEFYVDQV PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE VEILYNPKTKKHLGIAKVVFATVRGAKDAVQ HLHSTSVMGNIIHVELDTKGETRMRFYELLV TGRYTPQTLPVGELDAVSPIVNETLQLSDALK RLKDGGLSAGCGSGSSSVTPNSGGTPFSQDTA YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY SSRQPTPSYLFSQDPAVTFKARRHESKFTDAY NRHEHHYVHNSPAVTAVAGATAAFRGSSD LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG CLTQKSLFALFGGTEE\NMLLSVMAYDRFVAI CHPLYHSALMNPCFCAFLVLLSFFFLSLLDSQL HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF GGMEERHAPECDGL
1149	2499	Α	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE FRLVAADRSMGRYMLFGVINLICTGFLLMWC SSTINSIALTISYTYLTIFDLFSLMTCLISYWVTL RKPSPVYSFGFERLEVLAVFASTVLAQLGALF ILKESAERFLEQPEIHTGRLLVGTFVALCFNLF TMLSIRNKPFAYVSEAASTSWLQEHVADLSR SLCGIIPGLSSIFLPRMNPFVLIDLAGAFALCIT YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL GTNVSLRAA
1153	2503	A	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR PGNS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uctice	ŀ		314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Į į	•	j	l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	acquesto	/=possible nucleotide deletion, \=possible
			]	sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
						PT
1155	2505	Α	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	Α	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
1		ł	ł			EAAQRHCRASVSILRMRRPGQGSSRPARVPL
1157	2507	A	9327	152	292	RGPDSHRLREPPPSPP YERRGRSQGGSHPAGAQPGGRAIGAGWQS
1137	2507	^	9321	132	292	TERRORSQUOUSHFAUAQFOURAIDAG WQS   KEPLWEGLORSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
1		••	7520	•	130	LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
		ļ				VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
						SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
						RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
						LMEAGLPQKQAERADELFEAGLVTYVKLDER
1100	2610	<del></del>	0220			VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	A	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
		Ì				KRYVRILLLGEGAEHVADPVPGGRGVPRGEA DHTDQELREEIHKANVERVVHDVSQEATIEKI
]		ļ	j j			RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
		}				EAELPIMSQLTEIETCVEC
1161	2511	A	9341	1	390	NSRVDDFVAPGLSEAGKLLGLEFPERORLAA
				_		AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
1						DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
		ł	1			QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
						GELI
1162	2512	Α	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
1						TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
1						VVEPISDEDWYLFCGDTVEILEGKDAGKQGK VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
		i				GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
1				-		FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
						WIDGPKDTSVEDALERTYVPCLKTLQEEVME
						AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE
						G
1163	2513	A	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
		l			}	LPSSWAYRCASPHPDNFFVFLVESGFHHVGQ
						AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
1164	2514	A	9347	3	1099	FSSYQRNNPDLILNDTIMPNIK SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
''''	71 س		'דער	]	1077	HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
				İ	,	FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
						VAGASVGAGVWARNPRYRTEGEACVEFKA
]				, I		MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
			[ [	ĺ		VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
1 1						SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM
						SFLCLVVLYYTVWSLLFLRSLDVVAEQRRTH
1 1				i		VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
				ļ		YVSIFVPLWLSLLTLMATTFRRKGGNHWWF AIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKA
						LPLONKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
						VPEGVRLADGPGHCKGRVEVKHQNQWYTV
	-	ĺ				CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
				1		TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
						LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	Α	9363	201	387	PPILRWTPPSGKNFFFFFFESEFY/SSPRVECS
1167	2517	_	0269	707	1007	GAISAHLAHCNLCLPGSSDSPASAFQVAS
110/	2517	Α	9368	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
eotide seq-	rence sed- bebring		USSN 09/496	location correspondi	location corresponding to last amino	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
1		[		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	Joquan	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						PRPLILYAPAP\RPAGTAFIPHSHPPPPDLLRPT
		ĺ	i .			ATPA/TPCPSLPPPPRPLHPTQPSTALLPDPPPW
1168	2518	A	9375	511	15	PLPFPPPSS/RPPRPDCSTSYSPTFPPPT MMLSEETSAVRPQKQTRFNGAKLVWMLKGS
1	23.0	1	1373	311	13	PITVTSAVIIVLMLLMM/IFSPWLATHDPNAID
						LTARLLPPSAAHWFGTDEVGRDLFSRVLVGS
1						QQSILAGLVVVATTGMIGSPLECLFGELGGRA
						DAIFMRVMDIMRS/IPSLVLTMEKTAALGPSL
1169	2519	A	9377	42	410	FNAMQASSEH GNGRVAPRDPGAVASAEPGLTTHDSGVNPN
,	20.7	7.	,	42	410	NSARRMEAMASGSNWLSGVNVVLVMAYWS
			:			LVFVLLFIFAKRQIMRFAMKSLRGPHGPVGH
						NAPKDLKEEIDILLSRVHNIKYEP\HLLADDDA
1170	2520	A	9378	302	1303	GVSGFSASVLRQRRMEDELEPSLRPRTQIQGR
						ILLLTICAAGIGGTFQFGYNLSIINAPTLHIQEF TNETWQARTGEPLPDHLVLLMWSLIVSLYPL
				İ		GGLFGALLAGPLAITLGRKKSLL\VNNIFVVS
j j						AAILFGFSRKAGSFEMIMLGRLASWGVNAGV
						SMNIQP\MLPGGESAPKELRGAVAMSSAIFTA
						LGIVMGQVVGLSTTAATGLRGL\AGELEELEE
						ERAACQGCRARRPWELFQHRALRRQVTSLV
i						VLGSAMELCGNDSVYAYASSVFRKAGVPEA   KIQYAIIGTGSCELLTAVVSVSLEGALPPPAL
						WGGTPRSFALNQFTLQKKKK
1171	2521	Α	9381	2	412	RGPASAQEDERARTAPLERVRARGRMTTSSA
]						LFPSLLPCSWSTSNKYLAEFRAGKMSLKGTTE
						TPDKRKGLAY/IQQTDDSLIHFCWKDRTSGNV EDDLIIFPDDCEFKRLPQCPNGRVYVLKFKAG
[ [		ĺ				SKRLFFWMQEP
1172	2522	A	9384	20	355	GWNGRSTEASPAAEAPHVPHKETKAAMGTO
)		- 1				CTHGGKVRPDPHDMLTTVVHKIKLFVLCHSL
			1	-		LQLCAIMISDYLKSSIYTVEKRLGLFRPTSGLL
1173	2523	A	9393	430	87	ASFNEVGNTALIVLESY LCQCIVPGQQKETFSLNPSSATVRFYL*LSLO
		^		750	87	QRKEDQ*IIL*YHLNKDCLHIFMSAITLYMKI*
	}	j	J	J		KIFVLFDFNIMFETPFYII*FIFLFSQNLKRIRQV
<u> </u>						IRPPISFSKINNGP
1174	2524	A	9397	77	374	ERLEIGRLGGERGSGPASCLRVIDVSGMWDQ
	}	Į	l		ļ	RLVKLALLQLLRAFYGIKVKGVRVHRDCGTF ESSSTLIRVS*FGVPCNALAHFGVTHF*YILDF
<u> </u>		Į				LGML
1175	2525	Ā	9399	66	397	HESSRADRDKMDTRGSTYTDADPVNKSGGT
	Í	l	l			AKMNKWSKGKVRDKLNNLVLFDTATYDKL
	1		]		ļ	CKEVPNYKLITLAVVSERLKIPGSLARAALHE
1176	2526	A	9408	2	299	LLSRGLI*LVIQHIAQVIY  LDLTHVLSLSISLTVTILGTTFGMVIPLLDVVY
		••	7.30	-		GERGYAQNGDF*DAQLDDYSFSCYSHAQVN
	j	ľ			l	GAPNSLTRAYDDP*VKISGLECQKVGALVEV
1100						KCLNL
1177	2527	A	9416	2	402	CNFLRSSRIRVHSTPAASTMPPKVDPNEIKVV
		ŀ	.	ŀ		YLRCTGGEVRATSALAPKIGPLGLSSIKVGVD FV*ATGDWNVLIISVILTIRILLSHIFVVPPFFCF
			ł		}	DHLIAFWDLQSLIFLHVIFSLFITLLLFCFFSIF
1178	2528	A	9419	142	426	TPLFDLWPRVVLSWLETVLTSLRTRRAASGPP
						ACRIMPTTVDDVLEHGGEVHFLQKQMLYLL
1170	2520		0400	1450	1666	ALI*DTFAPIYVGIVFLGFTPDHRCRSPGVAEL
1179	2529	A	9420	1450	1655	LSSAGTKMNLN*KNYWPGASAHACNPSTLG GQSRCITRSGDRDHPG*HGETPSVLKIQKISRA
	İ		-		ĺ	WWRAP

CCCOTO	CEOR	3.600	Leco	Desdison	I business and	1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A
SEQ ID NO: of	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in		location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Ghutamine, R=Arginine, S=Serine,
1	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	)	l	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1180	2530	A	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
	1		1	1	ŀ	SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
		1				L
1181	2531	A	9436	2	274	PIAASLRMYNLQPYTEENLICTAFATMVETVP
			1			IARTILDRLTGIPHGYCFVE*ADWATADKCVH
1		ĺ			1	IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
1				-	- "	SMILK*MGAGDEKISAMGKARVDHRELYLGL
		1				LYPTEDYKLTFRARH
1183	2533	A	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR
1103	2333	^	7444	304	3	
					]	KGCSGWAPWLSLQCQHFGRPRWADHLRSGV
	İ		1			RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
			i			ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
			<u> </u>			ERG
1184	2534	Α	9462	391	655	LSGFKSLMPKIPLQYTYVRVRTTWSFCLPLDG
1			ľ			RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
			1			IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	Α	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
	1					WWWGWECWVRALKLSSGPAGPLACWVAK
						KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
	(	i				WSAVVQFWLTAASNSCFSLLSSWDYRCA
1186	2536	A	9468	275	452	HIPQLHTKTHYVPTRMVNKI*QIDNSKPWQR
11100	2000	1 ^	7400	213	432	
1187	2537	<del>  </del>	9469	200	2	GG*TGILTHCW*ESKLVQPLWKIVWHYQ
110/	2537	A	9409	388	3	EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ
		Į.				SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
		İ			ł	NPASPHPEAPQEPWDSASGSVGSFSLGRGAK
1	l					ASS*VPGKGRGPRQGSELLAETILELFLALAN
						S
1188	2538	Α	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
		1	1			GRLMANPEALKILSAITQPMVEEALAGLYRAC
						*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	A	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
		l l				PSLLKIQKISWAWWRAPVVPATWEAEAEEW
l		1				R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
		1				PROHCTLOTHRHLHPEAPVKV*KT*RLFPGLR
1						GASSCRRRRCNPVLAARKAGSPRSHSTRENC
						RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	Δ	9489	1	411	LADALCLSAAATGAVRPGARAQPSTRRRLSP
****		^	'*''	•	711	SVRVCCRAAAASNLLYSSCLQRHSERASEEG
		'	'			
					,	ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
1						MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ
1100	2542		0405	200	161	KEEELTAVNVK
1192	2542	A	9497	389	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
				•		CEEDERKMAREFLAEFMSTYVMMNIHMIVE
						KDTYSDHEEINTS
1193	2543	Α	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
						FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	A	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL
						LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
1 1			1			SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
						RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
''''	رجرے	^	2213		1444	
						AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC
						PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
, ,						LPRCSCAPLRSASAPQVS*CV*AVNLLPHNL*
						PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP
	j					APGSGPCGATARPSRGGRAGGSRARRPIPPGP
						GTRRTPSGCQNPAASGG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	- uuic		914	ng to first	acid residue	Q=Ghtamine, R=Arginine, S=Serine,
Cana	1	[	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	٠.			peptide	Schicuc	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1196	2546	A	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
11.50	2.540	<u> </u>	2210	223	100	AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
}		1	1		j	GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
1 ****	2,41	^	9321	209	****	HAMRHLTGNTSMAIRFL*ADSRFQVQRARYE
1					ł	APNWKYKYGY*IPVDMLC
1198	2548	A	9524	204	<del> </del>	
1170	2340	^	9324	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
						VKQEPNICHL*ETHFPFRDTYRLKEREQKKRK SSYS
1199	2549	A	9546	1785	1943	
1199	2349	A	y346	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1200	0.550	<del> </del>	0.540	102		V*QRGDGKNPGVTHLNRPVGTX
1200	2550	Α	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
1000	0.554		0.710			KAIYDKWTSDIMLNLQKL*AFFLRVIVRQI
1201	2551	Α	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
1						GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSQ
	ĺ	ŀ				YHVQHLATFIMDKSEAITSVDDAIRKLVQLSS
l		ł	1			KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
						PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
}						PDVHFFHCDEVEAELVHEYMESALTDCRLGK
						AMRP
1202	2552	Α	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
1						LDCERPPQGPLPSLPELAKTSYSDLTGLATED
1						*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
1						LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
						SKPRATPPLFCSLHTF
1203	2553	A	9568	517	738	RRKFERKQKQ*RYREGKQYRQRDKMKEWG
i l						EKEKRRREKGEREERKMRHRERKGESGQRD
						TMENWRVERLTEKER
1204	2554	Α	9573	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
ļ						DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
ĺ						EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
						HDCRHKEDAGVICSEFTALR '
1205	2555	. A	9577	64	424	ARGSCPTRPRTANGRMGETKDAPQMLVTFK
						DVAVTFFREEWRQLVLVHRTLYR*GMLETC
					,	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	Α	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
						SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
						NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
				ĺ		GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
						YNPGLPPLRTWNGQKLLWL
1207	2557	Α	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
]						PDGCRNVLRPKYYRLCDKAESWGIALETVPT
						GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
						THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
				l		FGILFSICFS
1208	2558	A	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
						FADAWADAW
1209	2559	Α	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
						GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
				į	l	RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
						MVCSIMM*FLLGITLLRSYMQSVWTRESQCT
			·		ľ	LLNASITETFNC
1210	2560	A	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
				- <b>-</b> -	-	DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
						KVQVKNNDLGLQATINNEANWIAHQDDFNW
1 1	ľ	1	· .	ľ	l	LLAELNTCQRQETADS***WSPKNSHVGKDS
						GELSAK
1211	2561	A	9620	316	610	QKHPGGGQLGRSPQEDSRFHNKASSGVSRVR
1						Z 300450101 ADDOM IN IN IN IN IN IN IN IN IN IN IN IN IN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L A-inc said converse (A-Alarica C-Constitution
						Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Ghrtamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	1		į.	peptide		/=possible nucleotide deletion, \=possible
1	1	l		sequence	i	nucleotide insertion
<b>}</b>	<del>                                     </del>	<del></del>		Sequence		
		l l				LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
-	1	l	i			GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE
Ł.,		L			l	TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR
		1				TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP
J	İ	J	1	}		AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
1			1			DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA
1213	2303	^	9024	2	330	
ł			į	Ì		TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ
		1	1	1		IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG
L	1		L		l	LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
						SHYCTPAWATEKDS
1215	2565	A	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
		١	7000		.20	EENFGEKLHDIGFGNGFLDKT*KAQATKAKI
1		1		1		DK
1216	2566	<del> </del>	000	201	76	I =.= ±
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
	1					RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI
1	Í		ĺ			NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
						LFGFILKAT
1217	2567	A	9655	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV
1.2	1 200.	١	1000	2000	2432	EEHHLQPVQVLQTLLHSATAGTGCRRPARPP
	ļ		1			
ĺ						PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
						GALGGRGGRALGGSRWPPPLPGETLFSGCKH
						RRRRGSDAAPGEEAGT
1218	2568	Α	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA
						VLMSAQEPWAIKEEHVIIQAEFYLNPDOSGEF
J	j	l	]			MLDFEGEDTFHGDMAKKETVWRLE*LARLD
	ì		<u> </u>			NFEAQRALANIAADQAALEIMDMGSDYTLIP
1		1	] .			NVPPKVTVL
1219	2569		9662		204	
1219	2309	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA
İ			i			YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ
1	i i	ĺ	1 1			LSQDPRNVWVFLATSGTLAGIMGMRFYHSG
			1			KL
1220	2570	A	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA
	1					PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN
1					·	GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF
1						PFKILSDAKLGTGPLRTFVCOVTSVIFYFTMYI
1						
1	i :					SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL
1222						K
1221	2571	A	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP
	1					QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV
1						VOILTALMSLSMGITMMCMASNTYGSNPISV
1	] .				ļ	YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG
		.				SLGMNITSS
1222	2572		9688	42	413	
1242	ا ۱۷۷۷	Α	7000	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF
•						PTDENIKRKWVLAMKRLDVNAAGIWEPKKG
			·	l		DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS
	<b> </b>					PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	Α	9696	308	564	RTSMGILYSEPICQAAYONDFGOVWRWVKE
						DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
	1 1			ay a		LKKECLCOPOKPERENLLALCCE
1224	2574		0700	<del>-,</del>	622	
1224	2574	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE
1						GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
[	( )			1		ELVERFALYGAIEQYNALDEYPAEDFTEVYLI
1 .			1	1	l	KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
				i		,
		ļ		1	1	PEFETVEETRKKLQMRKAYVVKTTENKDHY
			j		}	
				}		VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine O=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1225	2575	Α	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS TDDVHINDISLQDYIAGVRLILL
1226	2576	A	9713	82	492	QGLPSFLPAFGPSGSWLGPAPITGSSCNTVDT ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG ASVANKDIICYNLQAVGQIFYISSFLYTVNYI WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG LAAGKMNISIDLDTNYAELVLNVGRVTLGEN NRKKMKDCQLRKQQNENVSRAVCALLNSGG GVIKAEVENKGYSYKKDGIGLDLENSFSNML PFVPNFLDFMQNGNYF
1228	2578	Α	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN TLVLOKSDVEAVF
1229	2579	A	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP PLLEELGINFDHIWQKTLTVLHPLKVADGSIM NETDLAGPMVFCLAFGATLLLAGKIQFGYVY GISAIGCLGMFCLLNLMSMTGVSFGCVASVL GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG WCSFSASKIFISALAMEGQQLLVAYPCALLYG VFALISVF
1230	2580	A	9739	11	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG HFSPERPFMDYFDGVLMFVDISGKCKRDVCL MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS LRCGWSPAEELNYTVPGPGPAGEASPRQCRR YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC RDGWVYETPGSSIVTEFNLVCANSWMLDLFQ SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI NAAAGVLMAISPTYTWMLIFFLIQGLVSKAG WLIGYILITEFVGRRYRRTVGIFYQVAYTVGL LVLAGVAYALPHWRWLQFTVALPNFFFLLY YWCIPESPRWLISQNKNAEAMRIIKHIAKKNG KSLPASL
1232	2582	A	9753	164	517	PGPGMQGPPPTTPTSWSLPPWRAYVAAAVLC YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS FGILLWSGAGLSSSFISPRYSWLF
1233	2583	A	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM KRAYKSYVRALPLLKKMGINSILLRKSIGALE VACGIVMTLVPGRPKDVANFFLLLLVLAVLF FHQLV
1234	2584	A	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP CSRCGYGVYPAEKISCIDQIWHKACFHCEVC KMMLSVNNFVSHQKKPYCHAHNPKNNTFTS VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV FHWD
1235	2585	A	9767	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP CARPGHSARNNTDKNLPHTAIILVTSNTYTTI KINFQAGRSGSCL
1236	2586	Α	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		· ·	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		)		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						LERKQLNLEIYDPCSQTQKAKFSLTSELHWA DGFVIVYDISDRSSFAFAKALI
1237	2587	A	9793	266	515	NILAIIYFPFPRLFLLRDSOSNPKAFALTLCHH
						QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
						RYKPTRPVCITQFQGCS
1238	2588	Α	9802	537	967	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
						FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT
1 1						TDEDSTRFQIINEASKVPLLAEIYGIEGNIFRLK
		ĺ				INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
1220	2500		0005	105	710	GSLILADGKGDLKC
1239	2589	A	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
ì						MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
		ŀ				GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED   PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
						DIKFLTKGDNNEGDDRGSYK
1240	2590	A	9819	3	305	TDGRDPLPCAARRGGGGECCGAGWVAEWS
						POPLDPAMLLWMQGFVLEAVACQDNDDYLR
		l '	1			YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
						DPNSEEHG
1241	2591	Α	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
						LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
]						GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
1242	2502		0043		<b>400</b>	SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
						LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC DGSVKCTHQPCLIKWISERGCWSCELCYYKY
						HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
1			1			LFLIASISWLIWSTFSPSARWORODLLFOICYG
						MYGFMDVMIVAVDSEDMVQAAKEVGKRWS
						DIPP
1243	2593	A	9846	198	411	WRISHHAGKMPVMKGLLAPQNTFLDTIATRF
						DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
1044	0.504		00.40	11.5	7.0	FARTEVMQ
1244	2594	Α	9848	116	650	PICGFLYLCSAMASESSPLLAYRLIGEEGVAL
1						PANGAGGPGGASARKLSTFLGVVVPTVLSMF   SIVVFLRIGFVVGHAGLLOALAMLLVAYFILA
						LTVLSVCALATNGAVQGGGAYCILQHRWTG
						VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
						NVCGCAVSLLGLVESVLDVFGA
1245	2595	Α	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
						AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
						SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
] [		ļ				AMSLTEDVLAAALADHLPEDKWSAEKRRPL
		1		İ		KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
		1				YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
j İ						RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
j ľ	ĺ	ſ		ĺ	ĺ	SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
[ ]				l	ļ	EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
				l		LSGPTSEGLMTWELDRLLWARSVENLATATT
		- 1		1		TLTSLA
1246	2596	A	9850	114	464	PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS
						ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
	ŀ			ļ	ł	QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
<u> </u>						RRERPAGADSLSWGAGPRISSYV
1247	2597	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
. !		1		j		HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP GAILCSKHFQESDFESYGIRRKLKKGAVPSVS
	ł	. ]		ļ		LYKVFKYSSRCTS
			[			211.11.100NO10

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	peptide seq-	ĺ	USSN	location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	•••••		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	ŀ			peptide		/=possible nucleotide deletion, \=possible
	1			sequence	Į	nucleotide insertion
1248	2598	A	9853	58	444	RVDDFVYSKGGKDAGGADVSLACRRQSIPEE
						FRGITVVELIKKEGSTLGLTISGGTDKDGKPR
<b>!</b>		1				VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR
		1				LRHDEIITLLKNVGERVVLEVEYELPPPGGCP
1010	0.000		0000			WT
1249	2599	A	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV
1						RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA
						SGSGVAAGPAARHAPRRRCADAGEAVGASC
}						GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS
						GAGPAPGRAMRSTTLLALLALVLLYLVSGAL
						VFRALEQPHEQQAQRELGEVREKFLRAHPCV
						SDOELGLLIKEVADALGGGADPETNSTSNSSH
						SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK
						ELPHGGRCRETEGSQVAPRLPASPLCPGYGN
						VALRTDAGRLFCIFYALVGIPLFGILLAGVGD
1						RLGSSLRHGIGHIEAIFLKWHVPPELVRVLSA
						MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY
						FVIVTLTTVGFGDYVA
1250	2600	A	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF
						EWVYTDQPHTQRRKEILAKYPAIKALMRPDP
						RLKWAVLVLVQMLACWLVRGLAWRWLL
						FWAYAFGGCVNHSLTLAIHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH
						HRYLGGDGLDVDVPTRLEGWFFCTPARKLL
						WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV
						OLA
1251	2601	A	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR
						LESYRPOTOLSREDTGCNLQHISDRENIDDLN
						MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN
						HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI
						KCVALATYYHIKNRDPDGRMLLDIFDENLHPL
						SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL
	- 1					TAECAIVTLVYLERLLTYAEIDICPANWKRIV
						LGAILLASKVWDDQAVWNVDYCQILKDITVE
			1			DMNELERQFLELLQFNINVPSSVYAKYYFDL
1						RSLAEANNLSFPLEPLSRERAHKLEAISRLCED
1252	2602	A	9879	6	376	KYKDLRRSARKRSASADNLTLPRWSPAIIS KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT
'	2002	^	7013	ĭ	370	KAVGIKMGSLSTANVEFCLDVFKELNSNNIG
[ [		İ		l	•	DNIFFSSLSLLYALSMVLLGARGETEEQLEKV
						WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC
			1			NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA
						AMFGNC
1254	2604	A	9881	19	494	VISFQUITOTIMOSSTAHSPVFLVFPPEITASEYE
	Ì					STELSATTFSTQSPLQKLFARKMKILGTIQILF
					1	GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG
'	ļ		Į	ļ		SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA
1055	2605		0006			LGAIAGIILLTFEFHPRSKLHL
1255	2605	Α -	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT
		1	1	l		GATCVGLPNVGMCPQLSGALTFMYLQQGNQ
		ļ	l	1	ł	EATVAPDTMAQPYASAQFAPPQNGIPGEYTA
1256	2606	<del>_</del>	9902	05	200	PHPHPAPEYTGQTT
1230	2000	Α	3302	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIOEPGHAEGII GKROPKIKVEGEVEEDOGTK
	1	ł		I		DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH
		1		1		KEELE
1257	2607	A	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW
		<u></u> 1	7703			DIAMSTITUTO 1 17411 CHI DI LUCKUO W

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence.	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH QRRGPSCGASGDPQCVGSPHPQRARPLLARP GARLLPGHLPSPRPPRLPTGOPPAAAFRGPVR PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS YLGECGSSSYVTGAACISPVLRCREWFEAGLP WPYERGFLLHQKIALSRYATALEDTVDTSRL FRSRSLREFEEALFCHTKSFPISWDAYWDRND PLRDVDEAAVPVLCICSADDPVCGPPDHTLTT ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS HEVILESFRALTEFFRTEERIKGLSRHRASFLG GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL MAAAAGAAAAPGSREPQDRPECGAGHPGPR YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR ERPAARSGPEMRVRYPVVAAVLAPYLALSQD PMYKSSASGQGASGSYNHVREEMLIKAGGA MSRRVVRQSKFRHVFGQAAKADQAYEDIRV SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL PLAK
1259	2609	A	9919	693	935	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL PPRPGRSHRKRKLVSTK
1260	2610	A	9921	455	1082	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFQEALRNIGGQ IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE KKQKASQNLVVLAREDAGAEKIFRSNGVQLL QRLLDMGETDLMLAALRTLVGICSEHQSRTV ATLSILGTRRVVSILGVESQAVSLAACHLLQV MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKPRPTEGGPGAGSG RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF PTRVDHNGALLAFSPPFPQRQRRGTGATAES RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA PPRLPFCLQELQGRHALHTFSLERTCSYQDFL WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRPWEHRWSDPITLKMKGWG WLALLLGALLGTAWARRSQDLHCGACKAVR RRVROFNIYDY
1265	2615	A	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLTILLGLT EVAGEEELQMIQPEKLLLVTVGKTATLHCTV TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP RVTTVSDLTKRNNMDFSIRISSITPADVGTYY CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG FLSQVWWWLSSHPFMN
1266	2616	A	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC HPGYAGKTCNQGRKTV
1267	2617	A	10004	36	707	LPAPASTWSVARETMASSSVPPATVSAATAG PGPGFGFASKTKKKHFVQQKVKVFRAADPLV GVFLWGVAHSINELSQVPPPVMLLPDDFKAS SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP SQDELEHSLGESAAQGAAGVVLWVSWENTR

OPA TO	T CEC VE	1.37	Toro	I 70 - 41 - 1	T 10 - 10 - 10 - 10 - 10 - 10 - 10 - 10	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- eotide	peptide	1	in USSN		location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq- uence	İ		location	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi		M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
[		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł			İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	}	/=possible nucleotide deletion, \=possible
	↓	<del></del>	<del> </del> -	sequence	<u> </u>	nucleotide insertion
	<del> </del>	<u> </u>				TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHD
i	ì	i	ĺ			LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
ı			1			IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
		i	ı			AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
	1	<u> </u>				VQPCPICKEEFELRPQVFSIRG
1270	2620	Α	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
	1	ĺ				PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
1	ļ	İ				SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
		ĺ	]			RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
1	Í	l	1			QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
			1			QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
	l					DEIDV
1271	2621	Α	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
		l				FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
	ļ	1				LLGLAVGSYLVRRSRRPOVTLLDPNEKDLLR
						LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
1	1					SDEDOGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	A	10016	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
				-		SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
ļ	1					AQAAREPAAFWGPLARDTLVWDTPYHTVW
	i					DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
						PESVALIWERDEPGTEVRITYRELLETTCRLA
j	J	}				NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
	i	ł				CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
		l	!			ITFNQGLRGGRVVELKKIVDEAVKHCPTVOH
	J	1	1 1			VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
		1				VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
1		l	]			QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
1						WITGHSYVVYGPLCNGATSVLFESTPVYPNA
						GRYWETVERLKINQFYGAPTAVRLLLKYGD
			ļ			AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
						RVVGDSRCTLVDTWWQT
1274	2624	A	10017	1	3750	
12/7	2024	^	1001/	*	3130	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG
						KTLGSFFGSLPGFSSARNLVANAHSSARARPA
						ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
				İ		KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
						GVASVVDVAKGVVQGGLDTTRSALTGTKEV
1						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG
1	'			ĺ		TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
		!	1		i	LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
]					ļ	KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
		1		[	}	DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
						GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
						IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
i i				ĺ		KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
						LAKEAJQGGLDTTKSMVMGTKDTMSTGLTG
				ſ		AANVAKGAMQTGLNTTQNIATGTKDTVCSG
						VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
						SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
1				<u>.                                    </u>		AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
	i			j		TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
			1			VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
	1	ŀ	l	-		KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
				- 1	ļ	DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT
		ļ		į		GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
	ļ	ļ	j	j	ļ	QTGVDTAKTVLTGTKDTVTTGLMGAVNVAK
.						GTVQTSVDTTKTVLTGTKDTVCSGVTGAAN

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
		1100				
nucl	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	ŀ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	J	1.	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	Ì	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	(	ĺ	1	peptide	•	/=possible nucleotide deletion, \=possible
i				sequence		nucleotide insertion
<del></del>	<del>                                     </del>	<del>                                     </del>	<del> </del>	1 554 45550		VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA
i	İ	l				VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV
(	ſ	ŀ	1	1	<b>f</b>	
	1		i		•	TGAANVAKGAVQMGVDTAKTVLTGTKDTV
J	J	j	ļ	ļ	l	CSGVTGAANVAKGAVQTGLKTTQNIATGTK
1				1		NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT
			1	ļ		GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT
1			İ		l	VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT
ļ	)	1	l			AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG
1	i	1	]	]		VDASKAVLMGTKDTVFSGVTGAMSMAKGA
		i		l	1	VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA
1	ł			1		TGATHTGLSTFQNWLPSTPATSWGGLTSSRT
1	ļ	1	l	ŀ	]	TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP
		l	Į.	1	1	AWEAAATTKGLATDVATFTQGAAPGREDTG
[	[	ĺ	ſ		[	LLATTHGPEEAPRLAMLONELEGLGDIFHPM
ľ		1	1	(		NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD
1			1			LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL
1			1			
1000	2505	<u> </u>	10005	101		QDCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM
		ŀ			•	YKRKYSAANTKVEKKKKEKVLAPVTKPVGG
1	Ì	Į.				DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG
L		Ĺ				KKPFS
1276	2626	A	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS
			,			ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED
					· ·	EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL
	1		ſ	ĺ		LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK
ľ				1		CRIPNNSLPYFHKRPQARMLLLALFCVAVSV
				ĺ		VWGVFRNEDO
1277	2627	Α	10035	51	869	
12//	2027	Α.	10055	31	009	YSRFTVPLPATMASSEVARHLLFQSHMATKT
1						TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL
			i '			PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT
						LFNTNFEDYESSHFCPNVKLKAQTYELQESN
						VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN
(			1			PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE
				-		SSGQSAQQPYLPINSPPHRLADVADVHLFSSV
						LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS
						KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST
						VMGAVGESLSVOCRYEEKYKTFNKYWCROP
						CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL
•						TFTVTLENLTADDAGKYRCGIATILQEDGLSG
						EL BUDGEORONA NOS Y CCLENTANAS
1270	2620		10020	214	425	FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR
						SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA
						QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS
						FLRGLFGGNTRIEEACEMYTRAANMFKMAK
		- 1		' I	' i	NWSAAGNAFCQAAKLHMQLQSKHDSATSFV
	1	1		1	ì	DAGNAYKKADPQGKTARHVACYLCV
1281	2631	Ā	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG
1201	2051	^	10000	J20 .	010	
1292	2622	<del>-,</del> -l	10004	<del></del>	1640	PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY
		1				NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA
		i			ſ	ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF
	ľ	· · · · · · · · · · · · · · · · · · ·		ľ		YEFQLTAVSEGGVLSESSSTANITVVASDSPY
		1		ļ		GRFAFSHEQLRVSEAQRVNITHRSSGDFGHVR
		l				LWYKTMSGTAEAGLDFVPAAGELLFEAGEM
		Į				RKSLHVEILDDDYPEGPEEFSLTITKVELOGR
		ı		ł		GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN
	1	ļ	ļ			AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT
						CENTRAL PARTICIPATION A OPTIMITE A AVERGI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Ghutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGVRTII LTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM HLXRS
1284	2634	A .	10091	2	569	FVSPSRAMASALIYVSKFKSFVILVVTPLLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	A	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	Α	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1 .	478	MEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	A	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1		QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

NO- of NO- of NO- of NO- of NO- of No	SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Due coulde   Sed							
Dozation   Dozation			1				
Sequence   09/496   correspondition of the state of the			İ				
Unice     914   ng to first amino acid residue of peptide sequence							
mino acid recidule of peptide sequence purpose in the Threosine, V-Valine, W-Thyrophae, sequence peptide sequence peptide sequence per control of the sequence sequence sequence in the sequence		uaice	l				
residue of peptide sequence   y=Tyrosine, x=Unknown, **Siop codon, peptide sequence   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide insertion   y=Tyrosine, x=Unknown, **Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, peptide   y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, x=Unknown, y=Siop codon, y=Tyrosine, y=Tyrosine, x=Unknown, y=Siop codon, y=T	dence	l		714			
Population   Pop	ĺ	i	ĺ	1			
1293   2643   A   10124   2   989   PIMSLVRVVEFVAASSAQKTPSRLENYYMVC KADEKFNQLVFFRANSKORKELVLVK GVERMARDEVSTAVEVEFVAASSAQKTPSRLENYYMVC KADEKFNQLVFFRANSKORKELVLVK GVERMCHOGKMKYKRNKEMBEPKLQSGLUV CTDVMARGIDIPSVMAVLQVDPPSNASAFVH RCGRTARIGHGGSALVFLLVMESYNFLAN OKCPLOPEMKOPKTADLLPLLKSMALADRA VFEKGMKAFVSYVQAYAKHECNLEFULKDL DPASLARGEALLEVNKQKOPEMVALQVDPSNASAFVH RCGRTARIGHGGSALVFLLPMESYNFLAN OKCPLOPEMKOPKTADLLPLLKSMALADRA VFEKGMKAFVSYVQAYAKHECNLEFULKDL DPASLARGEALLRAWFAMPELRKQCOFPDVPV DVNTDTIPFKDKREKGORGKLLEGQRREKTEN EGRREFINKNASWSKOKAKKK   1294   2644   A   10129   91   1042   1047 TIMPSTANGLOPESYNFT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGASSISTICAG TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISTICAGCOSSUCCOPT TICCRATICAGCOSSISVIT TILCLAGCSLLAHASULVARGCOSSISTICAGCOSSUCCOPT TICCRATICAGCOSSICOLOGISTICATESCOSSICOLOGISTICATESCOSSICOLOGISTICATESCOSSICOLOGISTICATESCOSSICOLOGISTICATESCOSSICOLOGISTICATESCOSI	1		l	ĺ		sequence	
LGNVLTSTPAKTYNGKAESSDSGAESEEEE	ŀ		ſ	ĺ		[	
1293   2643   A   10124   2   989   PIMSLVRVVEFVAASSAQKTPSRLENYYMYC KADEKFNQLVHFLRNHKQEKILVFRYSSGAL CGGGRDSARMCSTCACVEYVGKALFULVK GVKINCHGKMKYKRNKFHEPKLQSGILV CTDVVAARGDIPEVNVLQVDPPSNASAFVH RCQRTARIGHGGSALVFLLPMEESYINFLAIN QKCPLQEMKPQRNTAQLLPKIKSMALADRA VPEKGMKAFVSTYVQAYAKHERULIPRIKADI. DFASLARGFALLRMFKMPELRGKQFDFPVYDNTDTTIPFRKINGEQRGKLLLOQGREKTEN EGRRFINKAWSKQKAKKK VKRAKKF GREKTENGKAWSKQKAKKK VKRAKKF VFEKGMKAFVSTYVQAYAKHERULIPRIKADI. DFASLARGFALLRMFKMPELRGKQFDFPVYDNTDTTPFRKINGEQRGKULLOQGREKTEN EGRRFINKAWSKQKAKKK VKRAKKF VFEKGMKAFVSTYVQAYAKHERULIPRIKADI. STANDAYSTILLIAFIFIMKRQDCSQMVNLPTQ LIFILLSVLLGIFGLAFAFELINQQTARAVSTILLIAFIFIMKRQDCSQMVNLPTQ GVIFALCFSCLLAHASNLVKLVRCCVSFSWT TILCIAGCSLQUILAFIELNQQTARAVSTILLIAFIFIXHATI TICIAGCSLQUILAFIFTVILIMTRGMMFVN MTFCQLNVDEVVLLYVYLFIMALTFFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMMFVN MTFCQLNVDEVVLLYVYLFIMALTFFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMMFVN MTFCQLNVDEVVLLYVYLFIMALTFFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTRGMFVN MTFCQLNVDEVVLLYVYLFIMALTFVSKAT FGGPCENWKQHGLIFTVTJEMTVMSMTL LAKAFTVTJEMTP FMGPCAPCHVA FASTANDA SAGRELADISTRACTENGCESTALATISTA VKRAMFGLGGLIGLNUKWGHGLIFTVTJEMTV VKRAMFGLGGLIGLNUKWGHGLIFTVTLENGCE TTCCRTTCCYSCCCYSCCCYCCCCCCCCCCCCCCCCCCCC	ļ				sequence		
1293	1						
KADEKPNQLVHFLRNHKOEKHLVFRYSSCALEVLVK	-	2512	Ļ.—	-			<u> </u>
CGRGRDSARMCSTCACVEYYGKALEVLVK   GVKINCHIGKMKYKRNKIMERIGENTALOGILV   CTDVMARGIDIPEVNVVLQVDPPSNASLYN   RCGRTARIGHGGSALVFILENEGSSALVFILENEGSSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGSALVFILENEGKGPDFVPV   DVRTDTIPFEDKIREKGRGKLLEQGREKTEN   EGRRKYKNKAWKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK	1293	2643	Α	10124	2	989	
GVKIMCHIGKMKYKRNIKIPMEERKLÖSGIL				1			
CTDVMARGIDIPEVNWULQYDPPSNASAFVH   RCGRARIGINGSALVFLIPMENTPLAIN   QKCPLQEMKPQRNTADILPKIKSMALAND   VFEKGMKAPKSVQAYAHENLIPKIKDIL   DFASILARGFALLRAPPKAPELRGKQFPDFVPV   DVNTDTIPFKDKIREKQRQKQLEQQRREKTEN   EGRRKFIKNKAWSKQAKKK   EGRRKFIKNKAWSKQAKKK   EGRRKFIKNKAWSKQAKKK   LIGIVVILLILAFILAFILMEKIQDCSQMVLPTQ   LLFLLSVLGIFGLAFAFILENQQTAPVKFIK   GVLFALGSGLLAHASNLVKLVRGCVSFSWT   TILCIAIGCSLQQIMATEYVTLIMTRGMMFVN   MTPCQLNVDFVVLLVVYLEMLFFYSKAT   FCGPCENWKQHGRLIPTIVLFSIIIVVVVSIIIL   LLRNPPGFQRQPOMDPVVCLALVTNAWVFL   LLYIVPELGLYRSCRQECPLQGNACPVTAV   HSRQVPNOGESIROKWKVLLNSDFSHSGA   STAN	1		i	1	1	}	
RCGRTARIGHGGSALVILHMEESYINFLAIN   QKCPLQEMKRQRNTADLIPKILSMALADRA   QKCPLQEMKRQRNTADLIPKILSMALADRA   QKCPLQEMKRAFYSYVQAYAKHECALIFRIKDI. DFASLARGEALIRJMEKMERJERGKOFPDFYVP   DYNTDTIPKDKIREKQROKULEQQRREKTEN   EGRKFIKNKAWSKQKAKKK   EGRKFIKNKAWSKQKAKKK   LIGIVYTILLLAFI-FIMRKIQDCSQWNVLPTQ   LLPLLSVLGIFGLAFAFIELHQQTAPWRYLFJ   GVLFALCFSCLLAHASNI.VKLVRGCVSFSWT   TILCIAIGCSLQUIIATEYYTLIMTGGMMFVN   MTPCQLNVDPFVVLLYVVLFIMALTFYVSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRLIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKQHGRIFTIVTYSKAT   FCGPCENWKVLLNSDFLSHSGA   RPRVVTTINSQCPGCPGSWRQV   HSFQVENQELSRDKWKVLLNSDFLSHSGA   RPRVVTTINSQCPGGSSWRQV   HSSQCPGGSSWRQV   HSSQCPGGSSWRQV   HSSQCPGGSSWRQV   HSSQCPGGSSWRQV   HSSQCPGGSSWRQV   HSSQCPGCWSCPGCQSDCGCQDLCQETCCRRSCCETTCCCHSXCC   FCGPCCWSCCRPCCQSVCCQPT   FCSRPSCCTTCCRTTCCRPSCCVSCCRCPQCCQSVCCQPT   FCSRPSCCTTCCRTTCCRPSCCVSCCRCPQCCQSVCCQPT   FCSRPSCCTTCCRTTCCRPSCCVSCCRPQCCQSVCCQPT   FCSRPSCCTTCCRTTCCRPSCCYQTTSCRSCCTTCCCTPSCCCPTTCCRTSCCSTTCCHSXCC   QEMYLRPDQTTRSSYRMSRILARRIQLVTK   QEGRAKASACMVRAGPT   ASQCENMERABRYSTANIARRIQLVTK   QEGRAKASACMVRAGPT   FSSELVDGGGRVSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIQLVTK   QEGRAKASCMVRAGRYSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIQLTVTK   QEGRAKASCMVRAGRYSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIQLTVTK   QEGRAKASCMVRAGRYSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIQLTVTK   QEGRAKASCMVRAGRYSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIQLTVTK   QEGRAKASCMVRAGRYSARFGGSSKAATVRSQPT   ASQCENMERABRYSTANIARRIARRIQLTVTK   QEGRAKASCMVRAGRYAGRYAGRYAGRYAGRYAGRYAGRYAGRYAGRYAGR					1		
	i .	ł		1	ł		
1294   2644   A   10129   91   1042   1047	1	j	ļ		ŀ		RCGRTARIGHGGSALVFLLPMEESYINFLAIN
DFASLARGFALLRMPKMPELRGKQFPDFVPV	1						
DFASLARGFALLRMPKMPELRGKQFPDFVPV	1	1	1			l	VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
		1				1	
	1	i	i	1		İ	DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN
ILGTVVTILLILAT FIJMRKIQDCSQWNVLPTQ   LIFLSVLGIFGLAFAFIELNQOTAPVRYFLF   GVLFALCFSCLIAHASNIVKLVRGCVSFSWT   TILCIAIGCSLLQIMATEYVTLIMTRGMMFVN   MTPCQLNVDFVVLLVYLMALTFFVSKAT   FCGPCENWKQHGRLFITTVLFSIIWVVWISML   LRGNPQFQRQPWDDPJVLAVTNAWVFL   LLYIVPELCILYRSCRQECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLLNSDFLSHSGA   LLYIVPELCILYRSCRQECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLLNSDFLSHSGA   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   GRGAPRQEGOSSWRQV   SWILDFPRGINSGQVGDLSFSQEKSLAQFRE   NIQDVLSALPNPDDYFLLRWLQARSFDLQKS   EDMIRKHMEFRKQQDLANILAWQPEVVRL   YNANGICGHDGGSSPVHVIGSQDFKGLLL   SASKQELLRDSFRSCELLLRECELQSKLGKR   VEKILAIFIGEGIGLRDKFGEILLQE   GVENCHT   TCCRTTCCRTSCVSSCCRPQCCQVCCQPT   CSRSCCQTTCCRTTCXPSSCCVSSCCRQCC   GVPVCCQFTCCRTSCVSSCCRQCC   GVPVCCQFTCCRTTCCRTTCXPSSCCVSSCCRQCC   GVPVCCQFTCCRTTCCRTTCXPSSCCVSSCCRQCC   GVPVCCQFTCCRTSCVSSCCRQCC   GVPVCCQFTCRTSSCCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPSCVSSCCRQCC   GVPVCCQFTCRTSPCCFTCCHPXCC   GSRRSCQTTCCHPXCC   GSRRSCQTTCCHPXCC   GSRRSCQTSPSCVSSCCRQCC   GSRRSCARSSQVSTAFASQCT   GSRSCAATVRSQFT   GSRRSCAATVRSQFT   GSRSCAATSAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVRSQFT   GSRSCAATVR				1			
LI_FILSV_GLEGLAFAFIIE_NQQT_APVRYPI_F	1294	2644	A	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA
LI_FILSV_GLEGLAFAFIIE_NQQT_APVRYPI_F							ILGIVVTILLLLAFLFLMRKIODCSOWNVLPTO
GVLFALCTSCLLAHASNLVKLVRGCVSSSWT   TILCIAIGCSLLQIILATEYVTLIMIRGIMFVN	i	ļ	l		į.		
TILCAIGCSLQIIATEYVTLIMTRGIMMFNN   MTPCQLNVDFVVLLVYVLFMALTFFVSKAT   FCGPCENWKQHGRLIFTVLFSIIWVVWISML   LRGNPCFCRQPQWDDPVCLAUTNAWVFL   LLYIVPELCL YRSCRQEPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLLNSDFLSHSGA   10133   376   518   RPRVVTHNSQWCFLPQDHFWLPGQSGAPG   GRQAPRQEGPGSSWRQV   HSFQVENQELSRDKWKVLLNSDFLSHSGA   RPRVVTHNSQWCFLPQDHFWLPGQSGAPG   GRQAPRQEGPGSSWRQV   GRQAPRQEGPGSSWRQV   HSFQVENDELSRDKWKVLLNSDFLSHSGA   RPRVVTHNSQWCFLPQDHFWGLPGQSGAPG   GRQAPRQEGPGSSWRQV   HSFQVENDELSRDKY   HSFQVENDELSRDK   HSF							
MTPCQLNVDFVULLVYVLELMALTFFVSKAT   FCGPCENWKQHGRLIFITVLFSIIWVVWISML   LRGNPQFQRQPQWDDFVVCIALVTNAWVFL   LLYIVPELCIL YRSCRQECPLQGNACPVTAYQ   HSFQVENQELSRDKWKVLINSDF.SHSGA   1975	1	1				1	
FCGPCENWKQHGRLIFITVLFSIIWVVWISML   LRGNPQFQRQPQWDDPVCIALVTNAWFL   LLYIVPELCILYRSCRQECPLQGNACPYTAYQ   HSFQVENQELSRDKWKVLLNSDFLSHSGA   RPRVVTINISQWCFIPQDHPGWLPGQSGAPG   GRGAPRQEGPGSSWRQV     1296	ĺ	[	·				,
LRGNPOF GROPQWDDPVVCIALVTNAWVFL				'			
LLYTYPELCILYRSCRQECPLQGNACPYTAYQ				ļ		1	
HSFQVENOELSRDKWKVLLNSDFLSHSGA				1	ł	}	
1295				•			
1296	1295	2645	Α	10133	376	518	
1296   2646   A   10135   3   551   EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICGHDGEGSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKHAIFGLEGLGJKDLWKPGIELLQE   1297   2647   A   10138   48   407   MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC   QPVCCQPTCCRPSCCETTCCHPXCC   QPVCCQPTCRPSCCETTCCHPXCC   QQPVCQPTCRPSCCETTCCHPXCC   QEMYLRFDQTTRSPYRMSRILARHQLVTKI   QQEIEAKEACDWLRAAGFQVAQLYEDSQFP   INIVAVKNDHDFLEKDLGEPLCRRINT   1299   2649   A   10161   1   393   PRFSELVDGRGVSARFGGSPSKAATVRSQPT   ASAQLEMBEAPKRVSLALQLPEHGSKDIGN   VPGNCSENPCQNGGTCVPGADAHSCDCGPGF   KGRCELACIKVSRPCTRLFSETKAFPVWEGG   VCHVV   VC	1.255	20.5	<b>'</b> `	10155	3,0	3.0	
NIQDVLSALPNPDDYFLLRWLQARSFDLQKS EDMLRKHMEFRKQQDLANILAWQPPEVVRL YNANGICDBGCSPVWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLLRECELQSQKLGKR VEKIIAIFGLEGLGRDLWKPGIELLQE  1297 2647 A 10138 48 407 MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCCPSCCETTCCHPXCC  1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSGFP NIVAVKNOHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEBAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGGK DPGVEGKHCEKKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTFV	1296	2646	A	10135	3	551	
EDMLRKHMEFRKQQDLANILAWQPPEVYRL YNANGICGHDGESPYWYHIVGSQDPKGLLL SASKQELLRDSFRSCELLIRECELQSQKLGKR VEKHAIFGLEGLGLRDLWKPGIELLQE  1297 2647 A 10138 48 407 MYSSCCGSVCSDQGCGQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVCCQPTCRPSCCETTCCHPXCC QPVLRPQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRINT 1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV 1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV 1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTFV	1	] =0.15	**	10100		55.	
YNANGICGHDGEGSPVWYHIVGSQDPKGLLL   SASKQELLRDSFRSCELLRECELQSQKLGKR   VEKILAFGELGLGLRDL WKPGIELLQE							
SASKQELLRDSFRSCELLLRECELQSQKLGKR							
VEKITÄIFGLEGLGLRDLWKPGIELLQE							
1297   2647   A   10138   48   407   MVSSCCGSVCSDQGCQDLCQETCCRPSCCE TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT CSRPSCCVTTCCRTSCCVSSCCRPQCC QPVCCQPTCCRTSCCTTCCHPXCC QPVCCQPTCCRPSCCETTCCHPXCC   1298   2648   A   10156   94   453   GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT   1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV   1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV   1301'   2651   A   10165   1   7545   PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQOPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV							
TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQFT CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQFTCCRTSCCTTCCTPXCC  1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRYSLALQLPEHGSKDJGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQOPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTFV	1297	2647	Α	10138	48	407	
CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC QPVCCQPTCCRPSCCETTCCHPXCC  1298 2648 A 10156 94 453 GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFFVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTFV							
1298   2648   A   10156   94   453   GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLFDQTTRRSPYMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDLGEPLCRRINT   1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV   1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMILAIADLLQVLSLPLRIFYYLKHDWPF VPV   1301   2651   A   10165   1   7545   PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1				1		
1298   2648   A   10156   94   453   GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT     1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEBAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV     1300   2650   A   10162   98   391   AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV     1301   2651   A   10165   1   7545   PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1						
QEMYLRFDQTTRRSPYRMSRILARHQLVTKI QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP INIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 1 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1298	2648	A	10156	94	453	
QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP NIVAVKNDHDFLEKDLGEPLCRRLNT  1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	,	20.0	1.	10150	74	,	
1299 2649 A 10161 I 393 PRFSELVDGRGRVSARFGGSPSKAATVRSQPT ASAQLENMEEAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			•				
1299   2649   A   10161   1   393   PRFSELVDGRGRVSARFGGSPSKAATVRSQPT							
ASAQLENMEBAPKRVSLALQLPEHGSKDIGN VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASILERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 1 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1299	2649	A	10161	1	303	
VPGNCSENPCQNGGTCVPGADAHSCDCGPGF KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			**	10101	•	3,3	
KGRRCELACIKVSRPCTRLFSETKAFPVWEGG VCHHV  1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	'						
VCHHV	] ]	·					•
1300 2650 A 10162 98 391 AKIASLERIMPANYTCTRPDGDNTDFRYFIYA VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301' 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	]			. 1		ĺ	
VTYTGILGPGLIGNILALWVFYGYMKETKRA VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301' 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHYPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEFGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2650		10162	OP.	201	
VIFMINLAIADLLQVLSLPLRIFYYLKHDWPF VPV  1301' 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEFGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1300	2030	Λ	10102	70	371	
1301' 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			-		1		
1301' 2651 A 10165 I 7545 PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			1				
KQLQCMPMEGRGRASSSISDLQGKGFEKGTG EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1201	2651	<del></del> _	10165	<del>-,</del> -	7545	·
EKHVPGVGSARHSPQASAGGSPWQRGKAQT RWLGKPDPGRKRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	1301	2001	A	10102	1	/343	
RWLGKPDPGRKRRRGSPQEEGGLRVSAAAR LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	*	İ					
LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV		ł	-				
PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	}	. ]		]			
LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	[	[	1	ĺ			
RTLSVEEPGVECNQLCLYADVTDPVLCLGQK DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV			i				
DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP ARRLSESLHVVDENKNESKIEREHKRRTSTPV	<b> </b>		1	ļ			
ARRLSESLHVVDENKNESKIEREHKRRTSTPV	[ . [	· 1	ľ	•	ĺ		
ARRLSESLHVVDENKNESKIEREHKRRTSTPV IMEGVQEETDTRDVKRQVERSEICTEEPQKQ	]	-					
	[	ĺ	ĺ		I		AKKLSESLHV VDENKNESKIEREHKRRTSTPV
	L						IMEGYGEEIDIKDYKKQVERSEICTEEPQKQ

NO: of No. of hed bed in DNO: insulation footion of the period of the property of the period of the period of the period of pe	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
soq- uence  ### unner			hod				
uence  ue			1				
uence    1914   maino seid residue of peptide pequence   peptide pequence   peptide seque							
mino acid reduce of peptide sequence se		dence					
residue of peptide sequence    Poptide in unclotide delicin, lypossible nuclotide delicin, lypossible nuclotide delicin, lypossible nuclotide delicin, lypossible nuclotide insertion    KSTILNSKHILKKDDSETPHLKSILKKEVKSS KEKPERREKTESEDKLSVKHKYKGDCMHKTTO DETELHSSEKGIK VERDINGKSOGOKIKSTOM FERKSKHRINERILSVILGKDGKPVSEVIIKTDE NYKERNNKKERILSAEKTKAEHKKRSSDSK (QKDSLGSK)GHOTLQRRSESYSEDKCDMDST NADSNLKPEVVHKEKRTKSLLEKLVLKS KSKTQGKQVKVFTELQEGATKQATTKKPD KKENTEENDSKEQKSKSVEDKFETGVEPV LETASSSAHSTQKDSSHRAKI, HAKEKYKSD KONSTIRTERKLSDGKKSSLKHSKSSKKDKKKD ENKSDDEDGGKVDSSHRAKI, HAKEKYKSD KONSTIRTERKLSDGKKSSLKHSKSDKKKD ENKSDDEDGGKVDSSHRAKI, HAKEKYKSD NOFTSTILERKLSDGKKSSLKHSKSDKKKD ENKSDDEDGGKVDSSHRAKI, HAKEKYKSD NOFTSTILERKLSDGKKSSLKHSKSSKKDKKKD ENKSDDEDGGKVDSSHRKARONSSLMEKKL SARLCERRGSLSQEAKGEKLAAATTIST SGOSSLORPKKSGDMTLIPGOPMEIDSEPGVVE NIFESKTOPNNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKTOPNNNSIGDDESEMKGKTS ATVQKDELRTCTADSKATAPAYKFGGTGV NIFESKATAPAYKFGGTGV	denoc			714			
peptide sequence    Possible nucleotide deteion, \possible   nucleotide insertion   RSTILKNEKHLKKDDSETPHILKSLILKKEVSS   KKPEREKTYPEGKU, SVHRKYKGDOMIKTO   DETELHSSEKGILKVERNIQKOGAVISTOM   TERKSHKINERLISVICKDGKPYSPETUKTIE   NVRKENNIKKERRISAGEKTA-GHIKSRESSIDSIN   QKDSIG-GSK-QHITUTGRSSSYSEDIKCDMDST    NVRKENNIKKERRISAGEKTA-GHIKSRESSIDSIN   QKDSIG-GSK-QHITUTGRSSSYSEDIKCDMDST    NMDSHLXPEEVVHKEKRITKSILLEGU, VLKS    SKKNTGKOVKVVETELDGBATKG-CHTYPYD    KEKNTEENDSEKQEKSKVEDKFFETOVEPV    LETASSSAHSTQKDSSHKARONSKILLEGU, VLKS    KSKNTGKOVKVVETELDGBATKG-CATTYPYD    KEKNTEENDSEKQEKSSKVEDKFFETOVEPV    LETASSSAHSTQKDSSHKARONSKILLEGU, VLKS    KSKNTGKOVKVVETELDGBATKG-CATTYPYD    KEKNTEENDSEKQEKSSKVEDKFFETOVEPV    LETASSSAHSTQKDSSHKARONSKILLEGU, VLKS    KSKNTGKOVKVVETELDGBATKG-CATTYPYD    KEKNTEENDSEKQEKSSKVEDKFFETOVEPV    LETASSSAHSTQKDSSHKARONSKILLEGU, VLKSK   KSKNTGKOVKVETELDGBATKG-CATTYPYD    KEKNTEENDSEKGEKSSLKHISKNDIKKKUL   SRRILCENPRIGGI-SQEMAK-GEFELA-ANT-SIT-ST-SSENGOR     KNESTELLEGU, KSKNTGK-CATTYPYD    KONTANTANEENVETOKAGKTON-SKARONSKILLEGU, VLKSK-CATTYPYD-SKARONSKILLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU, VLKSK-CATTYLLEGU							
Inteledide insertion	Ì					ooquanto	
KSTILKNEKHLIKKDDSEPIPILISELIKKEVSS REKPERKTYSEOKI SVIKHKYKGDOMIKITO DETELIHSSEKGIK VEENIOKOSQOTKLSSIDIK TERKSKRIRERIKLSIKORGKYVSEVIKTÜE NVREENINKERRILSAEKTKAEHKSRISSIDIS IQKOSLGSKOGHTILORGESSYSEDKOMDST NMDSNLKPEEVVHKEKRITKSILEEKLVIKS SKORGOVEVVETELOEGATKOATTREPD KEKNTEENIDSEKORKSKVEUKPEETETGVEPV LETASSSAHSTOKOSSHAALIP.ALKEVKSD KOSTISTILERILSDGHKSRSILKHSKODIKKKD DEKENTEENIDSEKORKSKVEUKPEETETGVEPV LETASSSAHSTOKOSSHAALIP.ALKEVKSD KOSTISTILERILSDGHKSRSILKHSKODIKKKD DEKENTEENIDSEKORKSKULHISKODIKKKD ENENDEVOKEVENISHKARONSSIMEKKI. SARILCENRIGSI.SQEMAKGEKKI.AANTILSTP SOSSLORPKKSGDMITLEGEDENIDEDEEPOVE NVFEVSKTODINRINNSIQDIDSENMKOKTS ATVOKODELRICTADSKATAPAYKERGTÖV NSINSEKHADHRSTLTIKKMHIOSAVSKONIDGE KEPIHERITEVINDISETVIRHILLSAPSENDRV QÖNILKINTAAEEHVAQGAATLEHSTILDSSENS LISSYTVYVILRESVIPPUPIPPURKTYLLEGSTA STSPADHSALINOSI.TVRESEVIKTSDSKEGG GEFTVYDTYARASTISKRRIIPEAHOLATILOKO GKYMPILGSKLATGVIVENENITEGOT. STSPADHSALINOSI.TVRESEVIKTSDSKEGG GEFTVYDTYARASTISKRRIIPEAHOLATILOKO GKYMPILGSKLATGVIVENENITEGOT. STSPADHSALINOSI.TVRESEVIKTSDSKEGG GEFTVTOTTAAASTISKRIIPEAHOLATILOKO GKYMPILGSKLATGVIVENENITEGOT. STSPADHSALINOSI.TVRESEVIKTSDSKEGG GEFTVTOTTAAASTISKATIFEAKOLATVILOKO GKYMPILGSKLATGVIVENENITEGOT. STSPADHSALINOSI.TVRESEVIKTSDSKEGG GEFTVTOTTAAASTISKATIFEAKOLATVILOKO GKYMPILGSKLATGVIVENENITEGOT. STSPADHSALINOSI.TVRESEOTIATILOKO GKYMPILGSKLATGVIVENENITEGOT. TOTAGARGROUNT TOTAGARGE GEFTVTOTLEGOTIATISTEVIKTOVILOKY AGAPATTISTEROBSKICAVASEEDAAVIL LAVIAVKERAVINSVYTEEKDDAVTSAGGEOTIVEN GARSTETILTSTREGESGICAVARSEEDAAVIL LAVIAVKERAVINSVYTEEKDDAVTSAGGEOTIVENE VEARAGAAMMANENINVDSMISGTEKGSKDT TYTICTGABGRSDIVEGTTTTISEVESDGAVTSAG TITAGGISSEEVOGSGNAMMINGRKKETIG TYTICTGABGRSDIVEGTTTTISEVESDGAVTSAG GGENTYTEERSEADDSIGGTAVST GEGAATTSTGTTEDGEGFASCT.OSEDSSEGFAIS SESEENGESAAMDSTAVTSESENDEDDITISTICAGG TYTICTGABGRSDAVITSTAGGRSTANDGTTVYGGC GPMTSAASDQSGLKVENTTISTICAGSS NOTAGGRSTATTSTAGGESGEOTIVENENING NEECDGLMATTASGGTTINTREPTOLDITIST GGGAATTSTOTTEDGEGFASCT.OSEDSSEGFAIS SESEENGEGAAMSTAAVSGGDOTTYTFOGCG GPMTSAASDQCAAVSEEDAANDENINGCCGOPV GGKEPOPLAASTEERGENAMPSLAGOG							
DETELHSSEKGIK VEENIQKOSQOYTK.SSDDIK TERSKERNERKIS/GKORGNESVSTEME NWRKENNIKKERRI.SAEKTRAEHKSRSSDISK (QKSIGS)GKOHGTURSESSYSEDKOLMDET NWRKENNIKKERRI.SAEKTRAEHKSRSSDISK (QKSIGSQHOHTU.GRSESSYSEDKOLMDET NWDSHLXPEEVVJKEKRRTKSI.LEEAL.VI,KS KSKTOGKOVKVVETELOGGATKQATTPKPD KERNTEENDSKORKSK VEDKPTEETGVEPV LETASSSAHSTOKDSHRAKUPLAKEKYKSD KOSTITELERI.SDGKRSSKLHSSKOLKKK ENKSDDKOEKVDSSHEKARGNSSLMEKKI. SRRICENGRGSLSQEMKSRSKLHSKOLIKKE ENKSDDKOEKVDSSHEKARGNSSLMEKKI. SRRICENGRGSLSQEMKOKSKIA. ANTISTT SOSSLORPKKSGDMTLIPEOEPMEIDSEFGVE NYFEVSKTODNENNISQDDISENMOKOKTS ATVOKDELRETCADSKATAPAYKPGRGTGV NINSEKHADHRSTLTKKMHIGASVSKOMPGE KEPHRIGTTEVNIDSETVHRMILSAPSENDRV QKINKINTAAEEBVAGODATLEISTINLDSSPS LISSVTVVPLRESVDPDVPLFDKRTVLEGSTA STSPADHSALPNOSLTVESSEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHGATLLDGKG GKVMPHGSKLTGVVENENITKEGGI.VDMA KKENDLNAEPNIKQTIKATVENGKKDGIAVD HVGLINTEKYAETVKLHKRSSPGKVINISD VERRIENSEVDTSAGSGSAPSVLHORNOGTE DVATOPPRAEKTSVATSTERGKDEVTLSPVK AGPATITSSETROSSEVALPETSIEADEGILIGT HERNIPLHVGAEASECTVFAAEEGGAVVTE GFAESETFLTSTREGESGEAVAESEDRAADL LAVTAVKLERAVNSVYTEERODAVTLSPVK AGPATITSSETROSSEVALPETSIEADEGILIGT HERNIPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTREGESGEAVASEDRAADL LAVTAVKLERAVNSVYTEERODAVTLSPVK AGPATITSSETROSSEVALPETSIEADEGILIGT HERNIPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTREGESGEAVASEDRAADL LAVTAVKLERAVNSVYTEERODAVTLSPVK AGPATITSSETROSSEVALPETSIEADEGILIGT HERNIPLTVGGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTREGESGEAVASEDRAADL LAVTAVKLERAVNSVYTEERODAVTLSPVK AGPATITSSETROSSEVALPETSIEADEGILIGT HERNIPLTVGGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTREGESGEGAVASEGGASGGGOSVINDGTE GESAVTSTGTGTEGGGGSGAFTGSEDSSEGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGTSTGTGTEEDEGGAVTSTGGAEGGAAVTSAG TEIRAGSISSEEDGASGGGGSGGSGGGSGGGGSGGGGGGGGGGGGGGG							KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
TERKSKHRINEKLISVLGKDGKPVSEVIIKTUR NVRKENNKKERISASKTKAEHKSRRSSDSK 10KDSLOSKOHGITLORRSESYSEDKCOMDIS NMDSNIKPEVVKKERRITKSLLEEKLVLKS KSKTOGKOVKVVETELQEGATKQATTEKPD KERNTENDEKGKOKSKVENDERPETGVEPV LETASSSAHSTOKDSSIRAKLPLAKEKYKSD KDSTSTELERKLSOHKSKSLEEKLVLKS KSKTOGKOVKVVETELQEGATKQATTEKPD KERNTENDEKGKOKSKVENDERPETGVEPV LETASSSAHSTOKDSSIRAKLPLAKEKYKSD KDSTSTELERKLSOHKSSKLEKKL ENKSDDKDOKEVDSSHEKARONSSLMEKKL ENKSDDKDOKEVDSSHEKARONSSLMEKKL SRILCEPRRGGISOSPMAGGEEKLAANTLSTP SGSSLORPKSSGDMAGGEEKLAANTLSTP SGSSLORPKSSGDMATLPEOPEMEDSEPVOV NVPEVSKTODINRNNSHQDDESDMACKTS ATVQKDELTCTADSKATAPAYKPGRGTOV NSNSEKHADHERSTIKKHRIQGAVSKMPGE KEPHRGTTEVADDSETVERMLSSASKMPGE KEPHRGTTEVADDSETVERMLSSASKEGG GEHTWOTDSETVERMLSSASKEGG GEFTYDTPAKASTISKRHIPPAGATALDEKGT QKNLKNTAAEEHVAQODATLEHSTNILDSSP SLSVTVVPLRESYDPVIPLDEKTVLBGSTA STSRADHSALPNQSLTVRESEVLKTSDSKEGG GEFTYDTTPAKASTISKRHIPPAGATLLDGKQ GKVMPLGSKLTGVVENENHTKEGGLVDMA KKENDLNAEPILKQTTLAVLHKHRSPGKVKDISID VERRHENSEVDTSAKSKHPAGATLLDGKQ GKVMPLGSKLTGVVENENHTKEGGLVDMA KKENDLNAEPILKGTTATVUNKKRGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRHENSEVDTSAGSGASPSVLHQRNOQTE DVATOPRRAEKTSVATSTEGKDKDVTLSVYK APPATTTSSETTGSEVALPCTSTEGADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGEAVAJECTSTEGDGGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGEAVADATSAGS KODOLSRDSEIVESTTTISTEVEDGAVTSAG TEIRAGSISSEEVDGSQGNMMMMGPKKETED TICTGAEGESDTVICSTTGGREERMTV GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESATTSTGTEDEGPASCTGSBEGFALS SESEENGESAMDSTVAKEGTNVPLVAAGPG GESATTSTGTEDEGPASCTGSBEGFALS SESEENGESAMDSTVAKEGTNVPLVAAGPG DICSSAKGIVESSVTSAVSGRDEVTPVPAGGC GMTSAASDQSDSQLEKVEDTTSTGLVGGG VDVLSGENESGVEVASGREERMTV GAGAADANDANENNVDSMGEGGDALSTTICAGS VDVLVSGENESGALGSTSTSOCOL DEGIVTSTGAKEEDGEGDEVTSTGRGNEGIG ASTCTGGAEGESGVLICESAGDORGITVSVV NECCOLARTTASAGGEGGPSSTTAEGGEGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSDSLEKVEDTTSTGLVGGG GPATSAASDGSBGGASSSTSSVCOCOL GEGVSSAV							
NVRKENNIKERRLSAEKTKAEHKSRRSISDS  () KDSILGSKOHGTIO, QRESESYSEBKCDMDST  NMDSNI.KPEVVHKERKRTKSLLEEKL.VLK.  KSKTOGKOVKVVETEL, QGGATKACHTEKPD  KERNTEENDESKORKSKVEDKPFETTGVETY  LETASSAHTOKOSSIRAALDH-AKELYKSPD  KERNTEENDESKORKSKVEDKPFETTGVETY  LETASSAHTOKOSSIRAALDH-AKELYKSPD  KERNTEENDENDEKDEKDOSSIRAACHAKELKSKDIKKAL  ENKSDENDENDEKDOSSIRAACHAKELKISTDIKEKL  BRALCENRGGI.SQEMAKGEEKLAANTI.STP  SGSSLQRPKKSODMTLEPGPMEDDSETOVETY  NYFEVSKTODNINNINSHODDISENMEQKITS  ATVQKDELRTCTADSKATAPAYKTOKGTOV  NSNSEKHADHRSTILTKKHHIQSAVSKMNPGE  KEPHRGTTEVMIDSETVHRMLSAFSENDSY  QKNILKNTAAEEHVAQODATLEHSTNLDSSPS  LSSVTVVPLRESYPPDVPLFDKRTVLESSEGG  EGFTVDTPAKASITSKRHIPPAHQATILLDGKQ  GKVIMPLGSKLTGVIVVENENTIKEGLOVMA  KKENDLNAEPALKOTKATVENGKEGIAVMA  KKENDLNAEPALKOTKATVENGKEGIAVMA  KKENDLNAEPALKOTKATVENGKEGIAVMA  KKENDLNAEPALKOTKATVENGKEGIAVMA  KKENDLNAEPALKOTKATVENKKKDGIAVVI  VERRHENSEVDTSAGSGSAPSVLHORNOGTE  DVATOPRAEKTSVATSTECKKDVILSID  VERRHENSEVDTSAGSGSAPSVLHORNOGTE  TOPTOPRAEKTSVATSTECKKDVILSID  VERRHENSEVDTSAGSGSAPSVLHORNOGTE  GFAESETFILTSKEGESGECAVAESEDRAADL  LAVHAYKEANVSVVTEENDATSAGSG  KCDGSLSRDSEIVGTTTTISEVESDA VTSAG  TEIRAGSISSEEVDGSQGNMMEMGPKKETED  TOTTCTGAGERSDIPVICSVTGAGREERMVT  GRAGVVLGDNDAPFGTSASGEDGSSGGAIS  SESSEENGESAMDSTVAKEGTNVPLVAAGPGE  GESAVTSTGTEDEGGPSACTGSEDSSEGAIS  SESSEENGESAMDSTVAKEGTNVPLVAACHGED  ASTCTGJGEESEGVLICESAEGDSOGGTVVVET  GRASTTSTGAEGEGGPSACTGSEDSSEGAIS  SESSEENGESAMDSTVAKEGTNVPLVAACHGED  ASTCTGJGEESEGVLICESAEGDSOGGTVVET  DICSSAKGIVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLYSGGVESSYTSAVSGRDETTPVGGCG  GYDTLASABDLAADLANGENNSPERSEDDILTTSICLVGGS  TYDLVSGGVESSYTSAVSGRDETTPVGGCG  GYDTLASABGRDSALADLAADLA  AISTSTTALCHERSASTRICHEENGLATAGGRNQAKETIT  EMMEGTRYTTEETFEAPPSAVSODSOLTAS  RSEEKDECAMISTSTREEPPSATTICAESAU  GOKEPPLASTRIKEERDGCALISTSIAEPCEGLASSY  GVVESENERAGTAMGAVLQAEBGRT  GKARPILATITUTTATAGGEST  ASYSAGRGLEG	1						DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
IQKDSLGSKQHGITLQRSESYSEDKCDMDS) NMDSNIKPEEVYREKRIKTSLEELLVLKS KSKTQGKQVKVVETELQEGATKQATTEKPD KEKNTENDESK GKRSKVEDERPETGVEPY LETASSSAHSTOKDSSIRAALPLAKELYXSB KDSTSTRLERKLSDGHKSRSLKHSKDIKKKD ENKSDDKDGKEVDSSIREAARONSSIMERKL SRRLCENRRGSLSQEMAKGEKLAANTLSTT SCSSLGPKKSGDMTLFFQEPREDSEFGVE NVFEVSKTQDNRNNNSHQDDSENMKQKT: ATVQKDELRCTADSKATAPAYRGRGTOV NSNSEKHADHRSTLTKKMHQSAVSKMPGE KEPHRGTTEVVIDSSTVARMLSSAVSKMPGE KEPHRGTTEVVIDSSTVARMLSSAVSKMPGE KEPHRGTTEVVIDSSTVARMLSSAVSKMPGE GKPKRSDTVPLRESPDVIPLDKSTVLGGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEG GFOTVDTPAKASTSKRHPEAHQATLLDGKQ GKVMPJGSKLTGVIVENENTKEGGLVDMA KKENDLMAEPHLAOTTALATURGK GKVMPJGSKLTGVIVENENTKEGGLVDMA KKENDLMAEPHLAOTTALATURGK GKVMPJGSKLTGVIVENENTKEGGLVDMA KKENDLMAEPHLAOTTALATURGK AGPATITSSETRQSEVALRCTNIEADEGLIGT DVATGPRAEKTSVATSTECKDKTVTLSPVK AGPATITSSETRQSEVALRCTNIEADEGLIGT HSRNPHLYGGASSECTVFAABEGGAVVTE GFASSETFLTSTKEGESGGCAVASESDRAADL LAVHAVKEANVSVVTEEKDDAVTSAGSEE KCDGSLSRDSEVVEGTITTIESVESDGAVTSAG TEIRAGSISSEEV DGSQGNMMRMGFKKETG TYTCTGGAGRSDNFVICSVTGAGPREERMVT GGGVVLGDNDAPGTSASGCGGASDSSCHATT GGAVATGTOTDEGGPASCTGSBDSSCHOTT GGSANTSTGTDGCGPASCTGSBDSSCHOTT GGSANTSTGTDGCGPASCTGSBDSSCHOTT GGSANTSTGTDGCGPASCTGSBDSSCHOTT GGSANTSTGTDGCGPASCTGSBDSSCHOTT GGSANTSTGTDGCGPASCTGSBDSSCHOTT GGSANTSTGTDCGCPASCTGSBDSSCHOTT GGSANTSTGTDCGCGAMSTRIGGCSGTANASCONSOLEKVEDTTISTGLVGGS YDVLVSGEVESVTSAVSGCDEVTTVPGGC GGMTSAASDCSDSQLEKVEDTTISTGLVGGS YDVLVSGEVESCVAANTSGCRANGKNOK VLISTSTTNDYTPQVSAITDVGGLSDAAGROD DICSSAKGTVESSVTSAVSGCDEVTTVPGGC GGMTSAASDCSDSQLEKVEDTTSTGLVGGS YDVLVSGEVESCVAANTSGECRANGKNOK VLISTSTTNDYTPQVSAITDVGGLSDAAGROD GGLATSTSKEEKDECAALISTSLAECCASVS GVVVSENERAGTQPVLISTATGFRAPPSAVD GGKEPGASTNSKEENDERONGTAADS SESCEDCAAMISTSIGEEFELPISSATTIKCAGS GYDVLVSGEWEGGTVATSTEGGUSTSSVEDC GGPYSSAVPQEEGDPSVTPAEMGDTAALIST SEGCGAAMISTGALGERINGVERAPSAP AASPLAASTSKEEKDECALISTSLAECCASVS GGVESDERAGTGVUSTSTARGGKNOGK VLISTSTTNDYTPQVSAITDVGGLSDAAGRONG VLISTSTTADCTRAASGCRANGTAL BENEGET GGGGGGSTALL HEARDANASCHALLGGGRANGTAR JESTAGCHAAMGCARGGRANGAL LILLIN HARRONGARGGGGGGGSTALLINGGARGGANGAL ALISTSTACCHAAMGCARGGGGGGGSTA							
NMDSNLKPÉVVHKERRRTKSLLEERLVIK KSKTOGKOVKVEPTELGGGATYKPD KEKNTERDASKORKSKVEDKPFETGYEVP LETASSAHTOKOSSIRAALPLAAELYKSD KEKNTERDASKORKSKVEDKPFETGYEVP LETASSAHTOKOSSIRAALPLAAELYKSD KOSTSTRLERKLSDOHKSRSLKHEKSKDIKKAL ENRSDAKOGKEVDSSHEKARONSSLMEKKL SRRLCENRRGISLOGEMAKGEEKLAANTLSTP SGSSLORPKKSODMTLEFGPMEDSETGYEV NVFEVSKTODNINNISHODDISENMIKKITS ATVQKDELRTCTADSKATAPAYKROKGTOV NSNSEKHADHRSTLTKKKHHOSAVSKMNPGE KEPHRGTTEVHIDSETYHRMLSAFSENDSY QKNLKNTAAELPHAQODATLEHSINLDSSPS LSSVTVVPLESYPPDVPLFDKRTVLESSENGG EGFTVDTPAKASITSKRHIPBAHQATLLIGGK GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPALKOTIKATVENOKKOGLAVAV HVGINTEKVAETVIKLKHKRSPGKVDISID VERRHENSEVDTSAGSGSAPSVLHORNOGTE DVATGPRAEKTSVATSTECKBVTLSFVK AGPATITSSETRGSEVALPCTSIEADEGLIGIG HSRNNPLHVGRAESCCTVFAAAESGCAVTE GRAESETFILTSKEGESGEAVAESGEDRAADL LAVHAYKEANVASVVTEEKDDAVTSAGSE KCDGSLSRDSEIVEGTITTISEVESDAVTSAGSE KCDGSLSRDSEIVERDAVATSAGSTAGSTAGSTAGSTAGSTAGSTAGSTAGSTAGSTA							
KSKTOGKÖVKVYETELGGATKOATTYRPD KERNTENDSEKQRESKVENEPFETGYEPY LETASSAHSTOKDSSHRAKI,PLAKEKYKSD KDSTSTRIERKI,SORIKKSSHEKARI,PLAKEKYKSD KDSTSTRIERKI,SORIKKSSHEKARINSSLMEKKI BNKSDDKDGKEVDSSHEKARGNSSLMEKKI SRRICENROSI,SQEMAKGEEKI,AANTI,STP SOSSI,QRYKKSGOMTLIPEGEHDSEPGVE NVFEVSKTQDNRNNNSRIQDIDSENMKQKTS ATVQKDELRICTADSKATAPYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPIHRGITEVINDESTVHRMILSAPSENDRY QKNI,KNTAAEEHVAQODATLEHSTNI,DSSPS LSSVTVVPLRSSTVPDVPHEJDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASTISKRHIPEDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG GGKVIMPLGSKLTGVIVVENENITREGG(JVMA KKENDI,NAEPNI,KQTIKATVHGKKXDGIAVD HVVGLNTEKYAETVALKJKKRSPGKVKDISD VERRNENSVDTSAGSGSAPSVLHQRNGQTE DVATOPRAERTSVATSTEDKORDVTLSPVK AGPATITSSETRQSEVALPCTSIRADBEGLIGT HSRNMPLHVQBASECTVFAAAEEGGAVVTE GFASESTILTSTKEGESGECAVAESEDRAADL LAVHAVKEANVSVVTEEKDAAVTSAGSE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEVEDGSQONNMRMGPKKTEG TVTCTGABGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSAGSGDGSVNDGTE GESAVTSTGTEDGEGPASCTGSEDSSEGAIS SESEENGESAMDSTVAKEGTSVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRCNEIGH ASTCTOLGEESEGVLICESAGDGGTVVEH VEAEAGAAMNANENNVTEEKSCHDDITTSVE NEECOGLAMTTASGDTTINDSLAGGKNQGK VLIJSTSTTNDYTTQVSAITDVEGGLSDALTTE ENNEGTRATTEEFFARPPSAVGGCTSVPLVAAGPC GGKRSAASDSVSVERCATUTTVELGAGSGGSTSSVEC GGMTSAASDSSVSLEKCHTITSTELVGGS YDVLVSGEVPECEVAHTSPSEKEDEDDITTSVE NEECOGLAMTTASGDTTINDSLAGGKNQGK VLIJSTSTTNDYTTQVSAITDVEGGLSDALTTE ENNEGTRATTEEFFARPPSAVGTSCORGVCVEH VEAEAGAAMNANENNVTERGERDGGLSDALTTE ENNEGTRATTEEFFARPPSAVGTSCORGVCVEH VEAEAGAAMNANENNVTERGERDGGGLSDALTTE ENNEGTRATTEEFFARPPSAVGTSCORGVCVEH VEAEGCAMTGSTGGEFFELPISSATTIKCAES LQPVAAAVEEEATGVLISTSLAGGGGSTSSVEC GGVSSAAVGGEGDAVTPAEAGGGGGTGSVEC GGVSSAAVGGEGDAVTPAEAGGGTGSTSVEC GGKPSAASGGRAVGGGAVALQDEDELTTEVELDLDA ALISTSTABCMPISALGRINGCCGPVR GGKEPGVLAVSTEEGHNOPSVHKSAGCGH PSAVCAKEEKHGKECPEIGPFAGRGGKRSTL HLINMEKNVLLINSLKKERDSPSVHKSAGCGAKA DSSVSSKTALAAVNTGAKKADDMPPQGGTAK DSSSKRYLAAVNTGAKADDMPPQGGTAK	1						
KEKNTEENDSEKQRKSKVEUKFPETGYEPY LETASSAHTSQKDSSHRAKLAKEKYKSD KDSTSTRIERKLSDGHKSPSILKHSSKJKKD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRRICENRROSLSQEMAKGEKIAANTLSTP SOSSLQRTKKSGDMTLIFEQEPMEIDSETGYE NVFEVSKTQDNINNINSIQIDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNSEKHADHRSTLTKKMIDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNSEKHADHRSTLTKKMIDSENDRV QKNILKITAAGEHVAQGDATUKLSDSKEGG KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNILKITAAGEHVAQGDATUKLSDSKEGG GEFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKRNDLNAEPNILKQTIKATVHGKKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKYKDISID VERRINENSEVITSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRGSSVLHQFRISGATVAGA GFATTSSTRIFTSGSSVLHQFRISGAD VERRINENSEVITSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSTRIFTSGSSVLHQFRISGAD VERRINENSEVITSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSTRIFTSGSSVLHQFRISGAD VERRINENSEVITSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSTRIFTSGSSVLHQFRISGAD VERRINENSEVITSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSTRIFTSGSSVLHQFRISGADATSAG GFACTSTRIFTSGSSTAGSTAGSEDANSGE KCDGSLSRDDSIVGSTTTISTSCDATSAG TEHAGSISSEVDGSQOGNMARMGPKKETEG TVTCTGABGRSDNFVICSVTGAGPREERMVT GAGVVLGDNAPPGTSAASGEGGSVNDGTE GESAVTSTGTEDGEGPASCTGSEDSSEGAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESESVTSAVSGKDSVTDVTFUGGGS TONLVSGEVPEGEVAHTSTSGROEIGH ASTCTGLGEESESVTSAVSGKDSVTPVPGGGC GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPEGEVAHTSTSGLSDGJGTTVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAGVIVESEVTSAVSGKDEVTPVPGGGC GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPEGEVAHTSTSGLSDGJGTTVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAGVPGEVEAVATSTGROSGGISTSVEDC GGYSSAAFVQEEGDPSVTPAEEMGDTAMISTS TSEGGCAVMIGGAVLQDEDRITTSUEDLSCASS GVVVSENERRATGTVMEKADGGGISTSSVEDC GGYSSAAPQEEGDPSVTPAEEMGDTAMISTS TSEGGCAVMIGGAVLQDEDRITTSUEDLSDA ALISTSTAECHPISASTIKEDLSCASS GVVVSENERRATGTVMEKADGGGISTSSVEDC GGYSSAAPQEEGPPSVRPSAQGGH PSAVCAKEEKHGKECPEIGPFAGROGKESTL HLINAEEKNVLLNSLKAERDSSPRITGAGGSST	1						I
LETASSAHSTQKDSHRALPILAKEKYKSD KDSTSTRIERALBGHKSSLKIKSSOMKKD ENKSDKDKCKEVDSSHRKARGNESIMEKKI SRELCENREGIS QEMAK GEEKLAANTLSTP SGSSLQRPKKSGDMTLIPEGEPMEIDSEPGVE NVFEVSKTQDNRNNSNGDIDSENKKQKTS ATVQKDELRTCTADSKATAPAYEPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPHRGTTEVNIDSETVHRMLISAPSENDRY QKNLKNTAAEEHVAQDATLEHSTNLDSSPS LSSVTVVELRESYDPDVPIPATLSSPSEDLRY QKNLKNTAAEEHVAQDATLEHSTNLDSSPS LSSVTVVELRESYDPDVPIPATLSSPSEDLRY QKNLKNTAAEEHVAQDATLEHSTNLDSSPS LSSVTVVELRESYDPDVPIPATLDGSTP STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLOSKLTGVTVENENTITEGGLVDMA KKENDLINAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYASTVALKHKRSPGKVKDISID VERRIPENSEVDTSAGSGSAPSVLHQRNOQTE DVATGPRRAEKTSVATSTEGHOKDVTLSPVK AGPATTSSETRQSEVALPCTSIEABEGLIGT HSRINPLHVQGAASSECTVEABEGGAVVTE GFAESETFLTSTKEGESGECAVAESDRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSETVEGTTTIFISEABEGGAVVTE GFAESETFLTSTKEGESGECAVAESDRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSETVEGTTTIFISEABEGGAVVTE GFAESETFLTSTKEGESGECAVAESDRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSETVEGTTTIFISEABEGGSVDTOTT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESATTSIGTEDGEGPASCTGSSESGGAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGBDVTTSTGRONEIGH ASTCTGLGESEGVLICESAEGDSGGTVVEH VEAAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVPPVGGCE GEMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTNSTEKEDEDITSVE NEECDGLMATTASGDTINONSLAGGKNOGK VLUSISTTNOYTPQVSAITDVEGGSDATATKCAES LQPVAAAVEERATGPVLISTAJEEGMPSAPP EAESPLASTISKEERDEEGLISTSTAEECEASVS GVVVSESNERAGTVMEEKADGGIBTSTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TEGCGAMMGRAVLQDEDERLTVEDLSDA AIISTSTAGCMALDGGGIBTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TEGCGAMMGRAVLQDEDERLTVEDLSDA AIISTSTAGCMALOPPOGGTVA DISSTENDERLENQLTADPEGN GDLSATEVIKHVPMSLABGROCGGBTTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TEGCGAMMGRAVLQDEDERLTVEDLSDA AIISTSTAECMPISASDRIBEENQLTADPPOGGTAA DSVSSIRVLAAAVNTGAKKADDMPPQGTVA DSVSSIRVLAAAVNTGAKKADDMPPQGTVAK DSSSIRVLAAAVNTGAKKADDMPPQGGTVA							
KDSTSTILERKLSDGHKSRSLKHSSKDHKKLD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRILCENRRGSLSQEMAKGEKLAANTLSTP SGSSLQPKKSGDMTLBEGPERDSERGVE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGKGTGV NSNSEKHADHRSTLTKKMHQSAYSKMMPGE KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEHVAQGDATEVHRMLLSAPSENDRV QKNLKNTAAEHVAQGDATEVHERMLLSAPSENDRV QKNLKNTAAEHVAQGDATEVLDGSTA STSPADHSALPNQSLTVRESENDRV QKNLKNTAAEHVAQGDATEVLDGSTA STSPADHSALPNQSLTVRESENDRV QKNLKNTAAEHVAQGDATEVLDGSTA STSPADHSALPNQSLTVRESENDRV QKNLKNTAGEGVENMA KENDLINAEPNLSQTTKATVENGKKDGLAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKKDGVTSDYK AGPATITSSETRSEVALPCTSIEADEGLLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETHLTSTKEGESGEGATSSEDASAGSGE KCDGSLSRDSETVGTTTTIESVESDAAVTSAG TERAGSISSEVDGSQOMMRMGPKKETEG TVTCTGABGRSDNPVLCSVTGAGFREEMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTEDGEGFASCTGSEDSSEGAIS SSSEENGESAMDSTVALSCHTNYPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRORGIGH ASTCTGLGEESEVVLESAGESQLGTVVPLGGS YDVLVSGEVVEGEVAHTSSEKEDEDUTTSTE QFAESEVALSTSTGAEEGEBLVTSTGAGKSKDT DICSSAKGTVESSVTSAVSGKDEVPVPGGGC GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVVEGEVAHTSSEKEDEDUTTSVE NEECDGLMATTASGDTNONSLAGGKNOGK VLHSTSTTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRYTTEEFFAPMPSAVGDDSQLTAS RSEEKDECAMSTISIGEEFELPISSATIKCAES LQPVAAAVEERATOPVLISTADEFGPMPSAPP EAESPLASTSKEEKDECALISTISIAECESAVS GVVVSENERAGTVPLEFFARMPSAVGDDSQLTAS RSEEKDECAMSTISIGEFEIPISSATIKCAES LQPVAAAVEERATOPVLISTADEFGFMPSAPP EAESPLASTSKEEKDECALISTISIAECESAVS GVVVSENERAGTVPLEFFARMPSAVGDDSQLTAS RSEEKDECAMSTSIGEFEIPISSATIKCAES LQPVAAAVEERATOPVLISTADEFGFMPSAPP EAESPLASTSKEEKDECALISTISIAECEASVS GVVVSENERAGTVTWEEKDAGGTSSVSCOC GGVSSANPQEGOPSVTPAEEMGDTAMSTS TSEGGEAVMGGAVLQDEDRITTEDLAGDROPPQGTVA ALISTSTACCHPISSASTIKEENOLTADEPGFMPSAPP EAESPLASTSKEEKDECALISTISIAECEASVS GVVVSENERAGTVTMEEKDSGTSSVEOC GGKEPGVLAVSTEEKANDEVHPSAGQGH PSAVCAKEEKHGKECFEIGPFAGROQKESTI. HLINAEEKNVLLINSLICKEDKSPSTETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEGTSGGTAK DSSVSSKINLAAVNISALKEDBSPTGTAGGSST							LETASSSAUSTOKUSSUBAKI DI AKEKAKSD
ENKSDDKDGKEVDSSIEKARGNSSLMEKKL SRRICENRGSLSQEMAKGERLAANTLSTP SGSSL ORPKKSGDMTLPEOEPMEIDSEPGVE NVFEVSKTQDNRNNNSHQIDISENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSMNPGE KEPIRRGTTEVNIDSETYPEMELISAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSYTVVPLRESYPDVPUFFRKTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHPAQATLLDGKG GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEFNLKQTIKATVENGKEGGLAVD HVVGLNTEKY AETVELKHRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAKTSVATSTGKDKDVTLSPVK AGPATTTSETRGSEVALPCTSIEADGLIIGT HSRNPPLHVQAAESECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRADL LAVHAVKIEANVASVVTEKKDAVTSAGSE KCDGLSRDSEIVBGTTTTSEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMRMROPKKETEG TVTCTGAEGRSDNPVICSTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAGCPCD DEGTVTSTGAKEGSDSVTVAGGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGGEPGSTVTSTGRORGKETGF GESAVTSTGITEDGGEPGEVVTSTGRORGHEIN SESEENGESGAMDSTVAKEGTNVPLYAGCPCD DEGTVTSTGAKEEDEEGGEVVTSTGRORGEGEAU SESEENGESGAMDSTVAKEGTNVPLYAGCPCD DEGTVSTGAKEEDEEGGEVVTSTGRORGEGEAU SESEENGESGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMGSTERGSKDT DICSSAKGTVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTSVE NEECDGLMATTASGDTINQNSLAGGNQGK VLUISTSTTINDTYPQVSATDVSGGLSDSALRTE ENMEGTRVTTEFFAAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGGMSDALATE ENMEGTRVTTEFFAAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAPP EASSPLASTSKEEKDEGSISTSSVEDC EGFVSSAVPQEEGDPSVTFAEEMGDTAMISTS TESEGCAVMIGGAVLQDEDTTITVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKIRKVPROSIDRITTITVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGHNGDSVHKPSAQGH PSAVCAEKEEKHGRECFEIGPPGRORESTI HILNAEENVLLINSLOCENDSPETGTAGGSST ASYSAGRGLEGMANSPAHLRCPEGTGGTAK DISSAKGRUALANDNPPVQGTVA	1						
SRRICENRGSISORMAGERIAANTI.STY SGSSLORPKSGOMTLIPEGEPMEIDSEPGVE NVFEVSKTQDNRNNISHQDIDSENMKQKTS ATVQKDELRTCTADSATAPAYKPGGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMMPGE KEPIHRGTTEVHIDSETVHRMLSAPSENDRV QKNLKNTAAEEHVAQGDTLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPPQSLTVRESVLKTDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMFLGSKLTGVVENENITKEGGLVDMA KKENDLNABEPILKQTIKATVENGKKDGJAVD HVVGLNTEKYAETVKLHKRSPGKVKDISID VERRINENSEVDTSAGSSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETTGSEVALPSTSTEABEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETHLTSTKEGESGCAVAESEDRADL LAVHAVKERANVSVVTEEKDDAVTSAGSEE KCDGSLSRDSETVESTTGSEVALPGRAVGSEE KCDGSLSRDSETVESTTTTISEVESDGAVTSAG TEIRAGSISSEVDGSGOMMARMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPOTASCEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SSESENGESAMDSTVAKEGTNVPLVAGPCD DEGITVSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSGIGTVVEH VEAAEGAAMNANENNVSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDOSDGLEKVTTSTGLYGGS VDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDVTPQVSATITDVEGGLSDALTRE ENMEGTRVTTEEFAPMPSAVSGDSQLTAS RSEEKDECAMISTSIGGEPISSATIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGFVSSAVPQEEGDPSVTTAEEMGDITAMIST TSEGCEAVMIGGAVLGPDRYSARCOGCESTSSVEDC GGRSAPPVLASTERGERINGPSVHKPSAGCH HINAEENVLLINSLOKENGERPOSVIERSAGCHL HLINAEENVLLINSLOKENGEPSPTTGTAGGSST ASYSAGRGLEGNANSPAHLRCPPEGTGGTSKA DSSVSSINYLLINSLOKENGENGENGCTS							
SGSSLQRPKKSGDMTLIPEQEPMEDSEPMEN NVFEVSKTQDNRNNISHQDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAYSMNPGE KEPIRKGTTEVNIDSETYERMLISAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPUFEKTYLEGSTA STSPADHSALPMQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLDGKQ GKVIMPLQSKLTGVTESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLDGKQ GKVIMPLQSKLTGVTESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLDGKQ GKVIMPLQSKLTGVTESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLDGKQ GKVIMPLQSKLTGVTESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLDGKQ GKVIMPLQSKLTGVTESITEADEGLIGT DVATOPRAEKTSVATSTEGKDKDVTLSPVX AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNINPLHVGARASECTVFAAAEEGGAVVTE GFASSTTTISTKEGESGGAVTAGAEEGGAVTE GFASSTTTISTKEGESGGAVTAGAEEGGAVTTE GFASSTTISTTSKEGSSEGVALAGEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEVEGTTTSEVESDGAVTSAGS TEIRAGSISSEEVDGSQQNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREEMVT GAGVVLGDNDAPPGTSASGEGGGSVNDOTE GESAVTSTGITEDCEGGPASCTGSEDSEGGFAIS SESEENGESAMDSTVAKSGTOVPLVAAGPCD DEGIVTSTGAKEEDEGEDVYTSTGROBIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VEAAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPPGGCE GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSSKEDDITSVE NEECDGLMATTASGOTINONSLAGGKNOGK VLIISTSTTNDYTPQVSAITDVEGGLSQLTAS RSEEKDECAMISTISGEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVYMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQEDERLTITRVEDLSDA AIISTSTTADCMPSAGNGENGLTAMSPERCH GGKEPGPVLAVSTEEGHNGPSVIREPSAGQGH PSAVCAKEKEEKHGKEPCPICTPAGGGSST ASYSAGRGLEGNANSPAHLRGPPGTFSGQTAK DISSNSSINYLAAAVNTGAIKGPCPTGPPR GGKEPGPVLAVSTEEGHNGPSVIREPSAGQCH HILDAEEKNVLINSLOKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPPGTFSQCTAK DISSNSSINYLAAAVNTGAIKADMPPVQGTVA							
NVFEVSKTQDNRNNNSIQDIDSENMKQKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPIHRGTTEVNIDSETVIRMLISAPSINDRY QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPLFPKRTVLBGSTA STSPADHSALPNQSLTVRESVLKTSDSKEGG GGFTVDTPAKASITSKRHIPPAHQATLLDGKG GKVMPLGSKLTGVIVENEVIKTSDSKEGG GGFTVDTPAKASITSKRHIPPAHQATLLDGKG GKVMPLGSKLTGVIVENEVIKTSDSKEGG GGFTVDTPAKASITSKRHIPPAHQATLLDGKG GKVMPLGSKLTGVIVENEVIKEGGLVDM AKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKVAETVKLKHRSPGVKVDISID VERRINENSEVDTSAGGSARSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTQSEVALFOSKDSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTGSEVALFOXABEDRAADL LAVHAVKIEANNVVTEEKDDAVTSAGSEE KCDGSLSRDSEITVETAAAEEGGAVVTE GFASSETTLTSTKEGEGGGAVTAG TERRAGSISSEVDGSQGMMRMMPKETEG TVTCTGAEGRSDNFVICSVTGAGPREEMVT GAGVVLGDDAPPGTSASCAGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTIGLGEESEGVLICESAEGDSQIGTVVEH VEAAEGAAMNANENNVDSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVPGGCE GPMSAASDOSDGLEKAMANANENNVDSMSGTEKGSKDT DICSSAKGTVESSVTSAVSGKDEVTPVFGGCE GPMSAASDOSDGLEKAMTSHGEFEPLISSATTIKCAES LOPVAAAVERATGPVLISTADFEGGLSDALATE ENMEGTRVTTEEFEAPMPSAVSGDSQLTAS RSEEKDECAMISTISGEFEPLISSATTIKCAES LOPVAAAVERATGPVLISTADFEGNSQLTAS RSEEKDECAMISTISGEFEPLISSATTIKCAES LOPVAAAVERATGPVLISTADFEGNFAMTS TSEGCEAVMIGGAVLQERUTHTVEDSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GCLSATEVSKIKEVPMPSLALANNCRCPGPVR GGKEPOPVLAVSTEEGHNGPSVIKEPSAGOGH PSAVCAEKEKHOKBCPEIGPPAGROGKESTL HLINAEENVLLNSLOKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPPGTTSGQTAK DSSVSSINYLAAANTGARNAPAHLRGPPGTSGQTAK DSSVSSINYLAAANTGARNAPAHLRGPPGTSGQTAK DSSSVSSINYLAAANTGARNAPAHLRGPPGTSGQTAK DSSSVSSINYLAAANTGANNPAHLRGPPGTSGQTAK		'					
NSNSEKHADHRSTLTKKMHIQSAVSKMAPGE KEPIHRGTTEVNIDSETVHRMLISAPSENDRV QKNIKNTAAEBHVAQQDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVIPLDKRTVLEGSTA STSPADHSALPNQSLTRSEVLKTSDSKEGG GGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEIPHLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSAPSVLHQNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGESGECAVASEDRAADL LAVHAVKIEANVNSVTEEDDDAVTSAGSEE KCDGSLSRDSEIVEGTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNEVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMGTEKGSKDT) DICSSAKGIVESSVTSAVSKGEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVEPCEVAHTSTSGLEDEIITSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLUSTSTTNDYTTQVSAITDVEGGLSDALRTE ENMEGTRYTTEFFEAPMPSAVSGDBSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFGPPMSAPP EASSPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSISTSSVEDC GGVSSAVPQEEGDPSVTPEEMGDTAMSTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNERGN GDLSATEVSKHVVPMETEEMGDTAMSTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENGLTADNERGN GDLSATEVSKHVVPMETEEMGDTAMSTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENGLTADNERGN GDLSATEVSKHVVPMETEEMGDTAMSTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENGLTADNERGN GDLSATEVSKHVVPMSLLAENNCCCPGPVR GGKEPGPVLAVSTEEGGINGFSVIKPSAGQGH PSAVCAKEKEKHGRECPEIGPPAGRGGKESTI. HLINAEEKNVLLNSLQKEDKSPTGTAGGSST ASYSAGRGLEGNANSPAHLRGPGTSGGTAK DSSVSISTRLAAVNITGAIKADDMPPVQGTVA	}						NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
KEPIHRGTTEVNIDSETVIRMILISAPSENDRY QKNLKNTAAEEHVAQGDATLEHSTNILDSSPS LSSVTVVPLRESVIPPVPLPURKTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATILLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTHAVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNGQTE  DVATOPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSTEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSE KCDGSLSRDSETVEGTITTISEVESDGAVTSAG TEIRAGSISSEVDGSQONMMRMOPKKETEG TVTCTGAEGRSDNPVCYSTGAGPREERMYT GAGVVLGDNDAPPGTSASQEGDSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGESSGVLICESAEGDSQITVVEH VEAEAGAAMMANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKCDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDTINONSLAGGKNOGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFGPPMPSAPP EASPLASTSKEEKDECALISTISIAEECEASVS GVVVSENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTTAEEMGDTAMISTS TEGCESAVMIGAVLQEDGRLTITHCDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATTSVSKHVPPAEEMGDTAMISTS TEGCESAVMIGAVLQEDGRLTITHVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATTSVSKHVPPAEEMGDTAMISTS TEGCCEAVMIGAVLQEDGRLTITHOLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHVVPPAEEMGDTAMISTS TEGCCEAVMIGAVLQEDGRLTITHVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHVVPSLEENPRAGGVKSTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASVSAGRGEGNANSPAHLRGPEGPFORFGRVSKTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASVSAGRGEGNANSPAHLRGPEGTSGGTAK DSSVSSINVLAAVNITGEGPPTGTAGGGST							•
OKNILKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTFAKASITSKRIIPEFAIQATLLDGKQ GKVIMPLGSKLI GTVVENENITKEGLVDMA KKENDLNAEPILKQTIKATVENGKKDGJAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRRAEKTSVATSTERDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNAVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASGEDGSVNDTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGDVTSTGRNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMNANENNYDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSGKDEVTPYGGCE GPMTSAASDQSDSQLEKVEDITTSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGRNQGK VLIISTSTTNDVTPQVSAITDVFGGLSDALTEE ENMEGTRVTTEEFEAPMPSAVSGDDSQLARTE ENMEGTRVTTEEFEAPMPSAVSGDSGLISTSSVEDC GGVSSAVPQEEGDFYTABEGMGTAMGSTATLERCCGSGTSTSSVEDC GGSSTAVPQEEGDFAGGSGTTANDSTATLRGPQTTSGGTTAK DSASAGRGLEGNANSPAHLRGPQTTSGGTAK DSASSAGRGLEGNANSPAHLRGPQTTSGGTTAK DSSASSGRGLEGNANSPAHLRGPQTTSGGTTAK DSSASGRGLEGNANSPAHLRGPQTTSGGTTAK DSSASGRGLEGNANSPAHLRGPQTTSGGTT							
STSYTVYPLRESYDPDVTPLFDKRTYLEGSTA STSPADHSALPNQSLTYRESEVLKTSDSKEGG EGFTYDTTAKASITSKRHIPEAHQATLLDGKQ GKYIMPLGSKLTGVIVENEINTKEGGLYDMA KKENDLNAEPHI, KQTIKATYENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGSGEGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVERGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNPVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDOTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAMMANDVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDITTSTGLVGGS YDVLVSGEVPECEVAHTSPSKEDEDITTSVE NEECDGLMATTASGDITNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELJSSATTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDSGISISSSVEDC EGPVSSAVPQEEGDPSVTPAEMGGTAMIST TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRTEEDNLTADNPEGN GDLSATEVSKIK VPMPSLLABNNCRCPGPVR GGKEPGPVLAVSTEEGHNORSYHKPSAQGGH PSAVCAKEEEHIGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLERGPGTTSGGTAK DSSVSSIRVLAAVMIGAIKADDMPPVQGTTAA							
STSPADHSALPNOSLTVRESEVLKTSISKEGG EGFTVDTPAKASITSKRHIPEAHQATILDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVGLMTEKYAETVKLKHKRSPGKVKDISID VERRENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRGSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSKEGSGECAVAESEGRAAVIE GFAESETFLTSKEGSGECAVAESEBRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGFKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPFSASGEGDGSVDDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESSENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVATISPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATPSVASGBOSQLENDETATS ENMEGTRVTTEEFEAPMPSAVSGDDSQLATS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATPSVASGBOSQLATTISCAES GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDFYTPAEEMGGTAMISTS TSEGCEAVMIGAVLQDERLTTTRVEDLSDA AIISTSTAECMPSVHKPSAGQGH PSAVCAAVEERATPGSVTHAEBMGDTAMISTS TSEGCEAVMIGAVLQDERLTTTRVEDLSDA AIISTSTAECMPSVHKPSAGQGH PSAVCAKEEEKHIGKECPEIGPFAGRGKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSFAHLRGPGTTSGGTAK DSSVSSIRVLAAVMIGAIKADDMPPVQGTTAA	1						
EGFTVDTPAKASITSKRHIPEAHQATLLDCKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPHLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLIKHRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITISSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEVDGSQGMMRMGMPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEOPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESSGVLICESAEGDSQIGTVVEH VEAEAGAAINMANENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIJISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSILEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTTRVEOLSDA AIISTSTAECMPISSURDHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGCECEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRILEGNANSPAHLRGPEQTTSGGTTAK DSSVSSIRYLAAVNTGALKADDMPPVQGTVA	l						
GKVIMPLOSKL TGVIVENENITKEGGLVDMA KKENDLNAEPNIK QTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHORNOGTE DVATGPRRAEKTSVATSTEGEKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLLIGT HSRINPLHVGAEASECTVFAAAEEGGAVYTE GFAESFTFLTSIKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVYTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSAGQEGDGSVNDOTE GESAVTSTGTIEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GFMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSKEDEDITTSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIJSTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTIGEFFLPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQBEDRITTITRVEDLSDA AIISTSTAECMPISASIDNHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHINGPSVHKTSAGQGH PSAVCAKEEKHGKECPEIGPPAGRGGKESTL HLINAEENVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGGTAK DSSVSSIRYLAAVNTGALKADDMPPVQGTVA	i !				ļ		
KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADBGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRADD. LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSETVEGTTITISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS SESEENDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQCH PSAVCAEKEEKHGKCPEIGPPAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRIEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	ľ i				İ		
VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETTQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETIFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEARAGAAMNAENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRITE ENNEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGFMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGFVSSAVPQEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQCTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	ļ					
DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITTSSETRQSEVALPCTSIEADEGLIGT HISRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESSEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASODITNQNSLAGGKNQGK VLIJISTSTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITTVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPSTTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRVLAAVNTGAIKADDMPPVQGTVA	i i			ľ		ĺ	HVVGLNTEKYAETVKLKHKRSPGKVKDISID
AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFASSETFLTSTKEGESGCAVAESEGRAVDL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGISSEVDGSQGMMRMGPKEFTEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEEJGGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGBUVTSTGRGNEIGH ASTCTGLGEESEGVLICESAAEGDSQIGTVVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATICPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTISAECECASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITTVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPTGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRVLAAVNTGAIKADDMPPVQGTVA				[			
HSRNNPLHVGAEASECTVFAAAEEGGAVYTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFISEVESDGAVTSAG TEIRAGSISSEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSECCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEFGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEERNVILINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSITYLAAVNTIGAIKADDMPVQGTVA							
GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNYSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSYTGAGPREERMVT GAGVVLGDNDAPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTITEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFLPISASTTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEFGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSSQTAK DSSVSSIRYLAAVNTGAIKADDMPVQGTVA				[			
LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITIFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNIFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTIPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDTINQNSLAGGKNQGK VLIISTSTTINDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGFSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							
KCDGSLSRDSEIVEGTTTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOFSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	i		ľ				
TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAECCEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTRAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKCPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA				ļ			
TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSLAEECEASVS GVVVSESNERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQGH PSAVCAEKEEKHGKECPEIGFFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ľ	f	i			
GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGCKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA							
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1		ł	1			
DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTFAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGGTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	]						
ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	1	ŀ	1	}		
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		.	-	ļ	. ]		
DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		}	ŀ	}	1	}	
GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	ļ	ŀ	.	_	
YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1	ŀ	ł	ŀ	1	j	-
NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1	ł		ł	1	1	
VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	j		j		J		
ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1	į	ł	ŀ	1	1	VLUSTSTINDYTPQVSAITDVEGGLSDALRTE
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			ļ	į	1		ENMEGTRYTTEEFEAPMPSAVSGDDSQLTAS
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1 1	ļ	1		1	}	1
GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ł	ļ	Ì		Ì	
EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ļ	Į	ŀ	1	}	
TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ļ		ļ	ļ	, 1	
AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	)		}	ŀ	]	j	
GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		l	1		l	Į	
GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		j	- 1	j	!	j	
PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			. [		ł		
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA						l	
ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	1	Ì	- 1	1	İ	Ì	HLINAEEKNVLLNSLQKEDKSPETGTAGGSST
			[	[	ſ		
EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP		-	ļ	}			
	L	1					EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, I=possible nucleotide deletion, V=possible nucleotide insertion  SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS SEENVCDIGNEESPLNVLGGLKLKANLKMEA YVPSEEKNGEILAPPESLCGGKPSGIAELQRE PLLVNESLNVENSGFRTNEEIHSESYNKGEISS GRKDNAFAISGHSVEADPKEVEEEERHMPKR KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC PETEPHATKEENSRDLEELPKTSSETNSTTSRV
1302	2652	A	10167	321	842	MEEKDEYSSSETTGEKPEQNDDDTIKSQE EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP EARETKCYVRSSVGCVEPLTTQAEVTENLDR KNSQQVFKLLKKK
1303	2653	A	10171	206	429	NMILLKKRRLLINSLGEGTINGLLDELLETNV LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITYI
1304	2654	A .	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF YHEAVVLFTQALKLNPQDHRLFGNRSFCHER LGQPAWALADAQVALTLRPGWPRGLFRLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH
1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT
1306	2656	A	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFTITVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR
1308	2658	A	10214	2	453	ECGGIRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNITQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM
1309	2659	A	10233	45	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV
1310	2660	A	10241	243	442	AFQLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK
1311	2661	A	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG
1312	2662	A	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide sequence	sequence	/=possible nucleotide deletion, \=possible nucleotide insertion
	,					LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF EHVCKPKVEHLIGYEVFECTHNMAYMLKKL LISYISIICVYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS KRFGVFLSEVSENKLREISLNHEWTFEKL
1313	2663	Α	10287	1221	266	GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAFDFARPL
1314	2664	A	10288	536	1890	LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLKVEFFPQEK
						DIKKEEITKEEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRGRSRSYSRSRSSWSKERLRERDRD RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN
	•					LPGMQPFPAQPPVVEGPPPPGLPPPPPILTPPPV NLRPPVPPPGPLPPSLPPVTGPPPPLPPLQPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTQKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPTPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIVLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

SEQ ID NO: of nucl- ectide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion  NEKLVDEKTILETSFHQHRERAEQLSQENEKL MNLLQERVKNEEPTTQEGKIELEQKCTGILE QGRFEREKLLNIQQQLTCSLRKVEEENQGAL EMIKRLKEENEKLNEFLELERHNNNMMAKTL
1318	2668	A	10303	333	879	EECRVTLEGLKMENGSLKSHLQG GECFIMAAVVQQNDLVFEFASNVMEDERQL GDPAIFPAVIVEHVPGADILNSYAGLACVEEP NDMITESSLDVAEEEIIDDDDDDITLTVEASCH DGDETIETIEAAEALLNMDSPGPMLDEKRINN NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ QVQEKYADSPGASSPEQPKRKKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLIIATALSL LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR LPRTLAGLLAGGALGLAGALMQTLTRNPLAD PGLLGVNAGASFAIVLGAALFGYSSAQEQLA MAFAGALVASLIVAFTGSQGGGQLSPVRLTL AGVXL
1320	2670	A	10323	441		KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV AVVDIQSDKAANVAQEINAEYGESMAYGFG ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI AKAAFISDFQLGDFDRSLQVNLVGYFLCARE FSRLMIRDGIQGRIIQINSKSDE
1321	2671	A	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD ILKCTLLVFGVRILYILKLNYTTEECDMKNMH YVDPDHVKRAQKYAQQVLQKESPPKFAKTS MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS ERKMRAHQVLTFLLLFVITSGASENASTSRGC GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL LGTLG
1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH QVLTFLLLFVITSVASENASTSRGCGLDLLPQ YVSLCDLDAIWGIVVEAAAGAGALITLLLMLI LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	A , .	10336		932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE NSVTHHEVKCQGKPLAGIYRKREEKRNAGN AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA PRKKAQGKTQQNRKLTIDFYPVRRSSRKSKAE LQSEERKRIDELIESGKEEGMKIDLIDGKGRG VIATKQFSRGDFVVEYHGDLIEITDAKKREAL YAQDPSTGCYMYYFQYLSKTYCVDATRETN RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL RGVTATFGRPAEWPGYLSHLCGRSAAMDLG PMRKSYRGDREAFEETHLTSLDPVKQFAAWF EEAVQCPDIGEANAMCLATCTRDGKPSARML LLKGFGKDGFRFFTNFESRKGKELDSNPFASL VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE QLYQDQEVPKPKSWGGYVLYPQVMEFWQG QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE DWLYERLAP
1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning mucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSD EDQGYVDLVIKVYLKGVHPKFPEGGKMSQY
						LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA ILKVPEDPTQCFLLFANQTEKDIILREDLEELQ ARYPNRFKLWFTLDHPPKDWAYSKGFVTAD MIREHLPAPGDDVLVLLCGPPPMVQLACHPN LDKLGYSQKMRFTY
1327	2677	Α	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR RYHRLSDMSMLAERRRKQKWAVDPQNTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIKVQVKNNHLGLGATINNEDNWIAHQ DDFNQLLAELNTCHGQETTDSSDKKEKKSFS LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL DCIFGKRQSKKTPEGDASPSTPEENETTTTSAF TIQEYFAKRMAALKNKPQVPVPGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA
1328	2678	A	10346	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE
1329	2679	Α	10351	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVITYIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKIISSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP
1330	2680		10352	34	2573	IPFLKSCCCCLFDFPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRILKVLDQVFYQRVSREGILSPSE LRKIPSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWPSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGLLKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISFFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE

	D Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of NO		ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1	tide	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq- uen		USSN 09/496	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
uence uen	CE	914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
acino.	1	7,7	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	- 1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ĺ		peptide	l code	/=possible nucleotide deletion, \=possible
			sequence		nucleotide insertion
1331 268	1 A	10353	1	2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEG
	Į.	-			AAGQQPTAPDKSKETNKTDNTEAPVTKIELLP
1	İ				SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE
	į	ļ			RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL
	l	1			SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP
1					IIMGANIGTSITNTIVALMQVGDRSEFRRAFA
1 1					GATVHDFFNWLSVLVLLPVEVATHYLEIITQL
	ļ				IVESFHFKNGEDAPDLLKVITKPFTKLIVQLDK
1 1	İ	i			KVISQIAMNDEKAKNKSLVKIWCKTFTNKTO
i l	- (				INVTVPSTANCTSPSLCWTDGIQNWTMKNVT
	j				YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV
1	1				LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP
		}			FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL
	l				TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL
	}	j			ASPGNALRSSLQIALCHFFFNISGILLWYPIPFT RLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIP
]	}				LTVFGLSLAGWRVLVGVGVPVVFIIILVLCLR
	i i				LLQSRCPRVLPKKLQNWNFLPLWMRSLKPW
		ļ			DAVVSKFTGCFQMRCCCCCRVCCRACCLLC
	- 1	ĺ			GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET
1333					FDNITISREAQGEVPASDSKTECTAL
1332 2682	2 A	10354	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPR
1	l				GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP
	ŀ				TSDLGEIHNWTELLDLFNHTLSECHVELSQST KRVVLFALYLAMFVVGLVENLLVICVNWRG
			[		SGRAGLMNLYILNMAIADLGIVLSLPVWMLE
] ]	1				VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF
j	j			1	LVCLSVDRYVTLTSASPSWQRYQHRVRRAM
	1				CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM
,	ļ			i	APFETYSTWALAVALSTTILGFLLPFPLITVFN
1 1	- 1	1 1	ĺ	1	VLTACRLRQPGQPKSRRHCLLLCAYVAVFV
1 1	ı	i i		1	MCWLPYHVTLLLLTLHGTHISLHCHLVHLLY FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL
	i				NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT
1	ł	! !			KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP
					TQPLTPS
1333 2683	A	10358	2	884	AAGAGADGREPASERASRAEPPAVAMGQND
		Į .	j	J	LMGTAEDFADQFLRVTKQYLPHVARLCLIST
]	}	] }	J	j	FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA
			ì		SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF
		'	ļ		GIIALQTIAYSILWDLKFLMRNLALGGGLLLL LAESRSEGKSMFAGVPTMRESSPKQYMQLGG
]		]	[	j	RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI
[ [			1	Ī	LVAIGFKTKLAALTLVVWLFAINVYFNAFWT
			ļ		IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL
		L			GPGGVSMDEKKKEW
1334 2684	A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFP
				l	ELPLPHVPGQESAKRRSARRFLIMSELTKELM
		}	į	1	ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS
] ]			ļ	1	ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS WRDCSQEEQKELLCHTLCDILESACCDHSGS
		j \$	ļ	ļ	YCLVSWLRGKTTEETASISGSPAESSCQVEHS
	İ		- {		SALAVEELGFERFHALIQKRSFRSLPELKDAV
				l	LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN
			ĺ	ļ	EIEDASEPLIDPVYGHGSQSLINLLLTGHAVSN
(			{	1	VWDGDRECSGMKLLGIHEQAAVGFLTLMEA
]					LRYCKVGSYLKISKIPYLDCLASETHLTVFFA
1 1	1		- 1	i	VDLAALMADE ADOEC ADDITECTION DECISION
			i	1	KDMALVAPEAPSEQARRVFQTYDPEDNGFIP DSLLEDVMKALDLVSDPEYINLMKNKLDPEG

1000-	l ore	136:	1 000	15.0		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl- cotide	peptide	1	in USSN	location	location	F-Phenylalanine, G-Glycine, H-Histidine,
seq-	seq-		09/496	correspondi	to last amino	Helsoleucine, Kelysine, Leleucine,
uence	ucute		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
1	1	1	1 -11-	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
j		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	Sequence	/=possible nucleotide deletion, \=possible
1	1	l l		sequence	1	nucleotide insertion
	<del>                                     </del>	<del>                                     </del>	<del> </del>	1	<del> </del>	LGULLGPFLQEFFPDQGSSGPESFTVYHYNGL
			1		1	KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
1		1			Į .	DTPIKRCLQTKWPYIELLWTTDRSPSLN
1335	2685	A	10375	82	2929	TRTKRRLGREKAMASPPRGWGCGELLLPFML
İ	ì	ĺ	İ			LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD
				1		LGLEPQELAERGVRIVSRGRTQLFALNPRSGS
l	İ	1	i	(	1	LVTAGRIDREELCAQSPLCVVNFNILVENKM
	[	[	İ			KIYGVEVEIIDINDNFPRFRDEELKVKVNENA
				•		AAGTRLVLPFARDADVGVNSLRSYQLSSNLH
-		1	1	i		FSLDVVSGTDGQKYPELVLEQPLDREKETVH
1	}	į	1		ļ	DLLLTALDGGDPVLSGTTHIRVTVLDANDNA
		1	1			PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE
		1	l		1	GINGKLTYSFRNEEEKISETFQLDSNLGEISTL
{		[		1		QSLDYEESRFYLMEVVAQDGGALVASAKVV
[	1	ĺ		(	ĺ	VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA
		1	i		İ	LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD
		l	ļ	1		NYYHLLTTRDLDREETSDYNITLTVMDHGTP PLSTESHIPLKVADVNDNPPNFPQASYSTSVT
						ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE
			}	}		DTFQGAPLSSYVSINSDTGVLYALRSFDYEOL
]		}	}	ļ	j	RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN
						DNTPEILYPALPTDGSTGVELAPRSAEPGYLV
		1				TKVVAVDKDSGQNAWLSYRLLKASEPGLFA
		ľ		l	1	VGLHTGEVRTARALLDRDALKQSLVVAVED
l		1		ì		HGQPPLSATFTVTVAVADRIPDILADLGSIKTP
	1					IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV
	· <b>,</b>	j	1		,	LRLRRWHKSRLLQAEGSRLAGVPASHFVGV
	]	]		·		DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY
		1				ADTLLSEESCEKSEPLLMSDKVDANKEERRV
						QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT
			l			WPNNQFDTEMLQAMILASASEAADGSSTLGG
		l	1			GAGTMGLSARYGPQFTLQHVLQGELGSDYR
					İ	QNYYIPGSNATLTNAAGKRDGKAPAGGNGN
1336	2686	A	10379	ī	557	KKKSGKKEKK RPRRROPSFSCRVLVLEDPPCFRFTNSMNOEK
	2000	**	103/3	1	""	LAKLQAQVRIGGKGTARRKKKVVHRTATAD
				·		DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI
						HFNNPKVQASLSANTFAITGHAEAKPITEMLP
			[	1		GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK
						PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	A	10380	1	1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL
						FGDLPEPERSPRPAAGKEAQKGPLLFDDLPPA
						SSTDSGSGGPLLFDDLPPASSGDSGSLATSISQ
						MVKTEGKGAKRKTSEEEKNGSEELVEKKVC
				[		KASSVIFGLKGYVAERKGEREEMQDAHVILN
						DITEECRPPSSLITRVSYFAVFDGHGGIRASKF
						AAQNLHQNLIRKFPKGDVISVEKTVKRCLLD
						TFKHTDEEFLKQASSQKPAWKDGSTATCVLA
. •						VDNILYIANLGDSRAILCRYNEESQKHAALSL
						SKEHNPTQYEERMRIQKAGGNVRDGRVLGV
.						LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND
				ļ	ſ	RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ
ĺ		{		ľ	İ	TREGKSAADARYEAACNRLANKAVQRGSAD
1338	2600		10305	<del></del>	500	NVTVMVVRIGH
1336	2688	Α	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG
	j	- 1		*	l	YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL KSIASADMDFNQLEAFLTAQTKKQGGITSDQ
						- n. s. m. s. m. i i i i i i i i i i i i i i i i i i
ļ		ļ				A VICKENIK CHRIKIDECI YWYOCHINDO
						AAVISKFWKSHKTKIRESLMNQSRWNSGLRG
				·		AAVISKFWKSHKTKIRESLMNQSRWNSGLRG LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alenine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \text{\text{\text{=}possible}} nucleotide insertion QPN
1339	2689	A	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG KCVICDSYVRPCTLVRICDECNYGSYQGRCVI CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL GSSKTDLFYERKKYGFKKR
1340	2690		10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK RPSWMVDNKRMRTASNFQWLLSTFILLYLM NQVNSQKKGAPHDLKCVTNNLQVWNCSWK APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL IPDTPEILNLSADPSTSTLYLKWNDRGSVFPHR SNVIWEIKVLRKESMELVKLVTHNTTLNGKD TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA LVGPRATSYTLVESFSGKYVRLKRAEAPTNES YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE NSSYLVALDKLNPYTLYTFRIRCSTETFWKW SKWSNKKQHLTTEASPSKGPDTWREWSSDG KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV KWEDIPVEELRGFLRGYLFYFGKGERDTSKM RVLESGRSDIKVKNITDISQKTLRIADLQGKTS YHLVLRAYTDGGVGPEKSMYVVTKENSVGL IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY PDIPNPENCKALQFQKSVCEGSSALKTLEMNP CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN RSDAKPENHVVESYCPPIIEEEIPNPAADETGG TAQVIYIDVQSMYQPQAKPEEQENDPVGGA GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI NSRQFLIPPKDEDSPKSNOGGWSFTNFFQNKP ND
1341	2691	A .	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL IKLRNLCLDWLQPETRTKEEIIELLVLEQYLTII PEKLKPWVRAKKPENCEKLVTLLENYKEMY QPEGESLHGVLVVSAGLRCPLGLSASTLLTW SGLDNSLSWAAVGMSCVLWDIELHHDFLGV ATKSVSTHAQGDAAQGLGGTTVRMWARDSN LATGVLLDDNNSDVTSDDDMTRNRRESSPPH SVHSFSGDRDWDRRGRSRDTEPRDRWSHTR NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FROGFRFNSTLVSRKRVLERKRRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS SLSSPSFTESQPIDFGAMPYVCDECGRSFSVIS EFVEHQIMHTRENLYEYGESFIHSVAVSEVQK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted and	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	S-15	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		Í	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1	i	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
[	ĺ	Í		peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			i — —	<del>                                     </del>	<b></b>	SOVGGKRFECKDCGETFNKSAALAEHRKIHA
ł	l	ł	l		i	RGYLVECKNOECEEAFMPSPTFSELOKIYGK
1					i	DKFYECRVCKETFLHSSALIEHOKIHFGDDKD
1	1		1	ĺ	1	NEREHERERERGETFRPSPALNEFQKMYG
i		1			ŀ	KEKMYECKVCGETFLHSSSLKEHQKIHTRGN
1	<b>i</b> :			ľ	! 	PFENKGKVCEETFIPGOSLKRROKTYNKEKLC
1	}	ļ		ļ		DFTDGRDAFMQSSELSEHQKIHSRKNLFEGR
1	l	ŀ				GYEKSVIHSGPFTESQKSHTITRPLESDEDEKA
		1	ļ	ļ	ļ	FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
				ĺ	ļ	ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN
1	}	ł		l		HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS
						HNGNELVESNEKGESSIYISDLNDKRQKIPAR
	1	l	1			ENPCEGGSKNRNYEDSVIOSVFRAKPOKSVP
			1			GEGSGEFKKDGEFSVPSSNVREYQKARAKKK
		[	1			YIEHRSNETSVIHSLPFGEOTFRPRGMLYECO
	]					ECGECFAHSSDLTEHQKIHDREKPSGSRNYE
						WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK
1						DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE
						NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
1			ł			KDDPDDKIYECEDCGLGFVDLTDLTDHOKVH
						SRKCLVDSREYTHSVIHTHSISEYQRDYTGEO
1			ł			LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM
					,	KGCDDGFIALLPMKPRRNRAAERNPALAGSA
1						IRCLLCGQGFIHSSALNEHMRLHREDDLLEOS
1						QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
	İ					ESFVNPAELADHVTVHKNEPYEYGSSYTHTS
						l
						FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
1 1						LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ
			,			VVLRIQGLNVEAAEPEVEAAEPEV
					•	EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE
1						QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE
i						GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
						HECTETFTSSTAFSEHLKTHASMIFEPANAFG
ł .						ECSGYIERASTSTGGANQADEKYFKCDVCGQ
1342	2600	A -	10702		1250	LFNDHLSLARHQNTHTG
1342	2692	Α	10393	2	1350	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
						ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS
			'			APAVLVVAVAVVVVVVSAVAWAMANYIHV
						PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
						HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN
	ŀ	ŀ				TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL
					- [	QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
						VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
				]		WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI
				ļ		LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII
} ]	ļ	ļ			ļ	YNEKQGDVQFIDYEYSGYNYLAYDIGNHFNE
		1	·		l	FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
)	j	ļ	j	J	ļ	EFKGFGTEVTEKEVEILFIQVNQFALASHFFW
		1	-			GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM
<u></u>						KPEVTALKVPE
1343	2693	Α	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA
	1	ĺ		ĺ		QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK
						TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN
	ł	ì	1			EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI
	1	ļ				LMYDITNEESFNAVQDWSTQIKTYSWDNAQ
		1		İ		VILVGNKCDMEDERVISTERGQHLGEQLGFE
1 1	J	1	1	ļ	ļ	FFETSAKDNINVKQTFERLVDIICDKMSESLET
		i				DPAITAAKQNTRLKETPPPPQPNCAC
1344	2694	Α	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS
L_ 1						LQQRTPAEMSPVLHFYVRPSGHEGAASGHTR

nucleotide sequence in ucleotide location corresponding uence uence uence uence uence location corresponding uence uence uence uence location corresponding uence uence uence location corresponding to last amino acid residue of peptide residue of peptide sequence location corresponding to last amino acid residue of peptide residue of peptide sequence location corresponding to last amino acid residue of peptide residue of peptide sequence location corresponding to last amino acid residue of peptide residue of peptide sequence location corresponding to last amino acid residue of peptide residue of peptide sequence location corresponding to last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide residue of peptide sequence location last amino acid residue of peptide residue of peptide residue of peptide location last amino acid residue of peptide residue of peptide residue of peptide residue of peptide residue of peptide residue of peptide residue of peptide	ine, an, on, ole TAEALPSA WLLPGSN ATGLGPV LLATLHDR
sequence  09/496  09/496  10 last amino acid residue of peptide sequence  14 last amino acid residue of peptide sequence  15 last amino acid residue of peptide sequence  16 last amino acid residue of peptide sequence  17 last amino acid residue of peptide sequence  18 last amino acid residue of peptide sequence  19 last amino acid residue of peptide sequence  28 last amino acid residue of peptide sequence  29 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  20 last amino acid residue of peptide sequence  21 last amino acid residue of peptide sequence  22 last amino acid residue of peptide sequence  23 last amino acid residue of peptide sequence  24 last amino acid residue of peptide sequence  25 last amino acid residue of peptide sequence  27 last amino acid residue of peptide sequence  28 last amino acid residue of peptide sequence  28 last amino acid residue of peptide sequence  28 l	TAEALPSA WLLPGSN ATGLGPV LATLHDR
uence  914  ng to first amino acid residue of peptide residue of peptide sequence  914  ng to first amino acid residue of peptide residue of peptide sequence  914  ng to first amino acid residue of peptide sequence  92-Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptopha y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion  914  New Tryptopha y=Tyrosine, X=Unknown, *=Stop cod /=possible nucleotide deletion, \=possible nucleotide insertion  915  RKLQGKLPELQGVETELCYNVNW EETKKLMWLFGCPLLLDDVARESV DLLLEVGPRLNFSTPTSTNIVSVCR DRVETTRRYRLSFAHPPSAEVEAIA	TAEALPSA WLLPGSN ATGLGPV LLATLHDR
amino acid residue of peptide residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop cod peptide sequence   T=Threonine, V=Valine, W=Tryptophi Y=Tyrosine, X=Unknown, *=Stop cod peptide sequence   T=Threonine, V=Valine, W=Tryptophi Y=Tyrosine, X=Unknown, *=Stop cod peptide   Possible nucleotide deletion, \=possible nucleotide insertion   RKLQGKLPELQGVETELCYNVNW   EETKKLMWLFGCPLLLDDVARESV   DLLLEVGPRLNFSTPTSTNIVSVCR   DRVETTRRYRLSFAHPPSAEVEAIA	TAEALPSA WLLPGSN ATGLGPV LLATLHDR
residue of peptide peptide sequence Y=Tyrosine, X=Unknown, *=Stop cod peptide sequence requence residue of peptide sequence requence requence residue of peptide sequence requ	TAEALPSA WLLPGSN ATGLGPV LLATLHDR
peptide /≔possible nucleotide deletion, \≔possible sequence nucleotide insertion  RKLQGKLPELQGVETELCYNVNW  EETKKLMWLFGCPLLLDDVARESV  DLLLEVGPRLNFSTPTSTNIVSVCR  DRVETTRRYRLSFAHPPSAEVEAIA	TAEALPSA WLLPGSN ATGLGPV LLATLHDR
sequence aucleotide insertion  RKLQGKLPELQGVETELCYNVNW  EETKKLMWLFGCPLLLDDVARESV  DLLLEVGPRLNFSTPTSTNIVSVCR  DRVETTRRYRLSFAHPPSAEVEAIA	TAEALPSA WLLPGSN ATGLGPV LLATLHDR
EETKKLMWLFGCPLLLDDVARESV DLLLEVGPRLNFSTPTSTNIVSVCR DRVETTRRYRLSFAHPPSAEVEAIA	WLLPGSN ATGLGPV LLATLHDR
DLLLEVGPRLNFSTPTSTNIVSVCR. DRVETTRRYRLSFAHPPSAEVEAIA	ATGLGPV LATLHDR
DRVETTRRYRLSFAHPPSAEVEAIA	LATLHDR
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	
MTEQHFPHPIQSFSPESMPEPLNGP	ו מממא חומו
LALEKANQELGLALDSWDLDFYTI	
NPSTVEAFDLAQSNSEHSRHWFFK QKLVHSLFESIMSTQESSNPNNVLK	
IQGKEVRFLRPEDPTRPSRFQQQQQ	
AETHNFPTGVCPFSGATTGTGGGRIF	
RGAHVVAGTAGYCFGNLHIPGYNI	, ,
QYPGNFARPLEVAIEASNGASDYG	
LAGFARSLGLQLPDGQRREWIKPIN	MFSGGIGS
MEADHISKEAPEPGMEVVKVGGPV	
GAASSVQVQGDNTSDLDFGAVQR	
KMNRVIRACVEAPKGNPICSLHDQ	
NVLKELSDPAGAIIYTSRFQLGDPT GAEYQESNALLLRSPNRDFLTHVS	
CFVGTTTGDRRIVLVDDRECPVRRN	
PTPPPTPVDLELEWVLGKMPRKEF	
MLQPLALPPGLSVHQALERVLRLP	
LTNKVDRSVGGLVAQQQCVGPLO	
VVALSHEELIGAATALĞEQPVKSLİ	
RLAVAEALTNLVFALVTDLRDVKC	CSGNWM
WAAKLPGEGAALADACEAMVAVI	
VDGGKDSLSMAARVGTETVRAPG	
VCPDITATVTPDLKHPEGRGHLLYV	•
HRLGGTALAQCFSQLGEHPPDLDL FSITQGLLKDRLLCSGHDVSDGGLV	
AFAGNCGLQVDVPVPRVDVLSVLF	
LEVQEPDLAQVLKRYRDAGLHCLE	
AGPHAMVRVSVNGAVVLEEPVGE	
TSFQLDRLQAEPRCVAEEERGLREF	
LPPTFPKASVPREPGGPSPRVAILRE	EEGSNGDR
EMADAFHLAGFEVWDVTMQDLCS	GAIGLDT
FRGVAFVGGFSYADVLGSAKGWA	
RAGAELRRFRKRPDTFSLGVCNGC	
WVGGDPNEDAAEMGPDSQPARPG	_ 1
SGRYESRWASVRVGPGPALMLRGI   VWSAHGEGYVAFSSPELQAQIEAR	
ADDDGNPTEQYPLNPNGSPGGVAC	
HLAVMPHPERAVRPWQWAWRPPP	
PWLQLFINARNWTLEGSC	
1345 2695 A 10396 65 642 GVRGFWAGTMASRAGPRAAGTDG	SDFQHRE
RVAMHYQMSVTLKYEIKKLIYVHL	
AKMSVGHLRLLSHDQVAMPYQWE	
LPSLLGLLSFPRNNISYLVLSMISMO	
YGSMEMFPAAQQLYRHGKAYRFLI	
MYLVLVLAVQVHAWQLYYSKKLL	DSWFTST
QEKKHK   1346   2696   A   10398   1   718   DDFVRCGPQSAAMGASARLLRAVI	MGARGE
1546 2696 A 10598 I 718 DDFVRCGPQSAAMGASARLLRAVI GKGTVSSRITTHFELKHLSSGDLLRI	
EIGVLAKAFIDOGKLIPDDVMTRLA	
TOYSWLLDGFPRTLPOAEALDRAY	
NVPFEVIKQRLTARWIHPASGRVYY	`
TVGIDDLTGEPLIQREDDKPETVIKF	LKAYED.
QTKPVLEYYQKKGVLETFSGTETNI	KIWPYVY
AFLQTKVPQRSQKASVTP	
1347 2697 A 10402 153 1969 KHRQENNALDMAPEIHMTGPMCLI	
VANPEALKILSAITQPVVVVAIVGLY	IKIGKSY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}		j		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	j	ł	peptide	•	/=possible nucleotide deletion, \=possible
}	}	j		sequence	ļ	nucleotide insertion
			i			LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
1	j	ļ	İ		ļ	PHPKKPEHTLVLLDTEGLGDVKKGDNQNDS
	<u> </u>					WIFTLAVLLSSTLVYNSMGTTNQQAMDQLYY
				j		VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
	[			1	1	WTLRDFSLDLEADGQPLTPDEYLEYSLKLTQ
	Ì	l '		1		GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
1	[			ĺ	1	HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
1	Ì			ŀ		FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
				ĺ		GDLPCMENAVLALAQIENSAAVQKAIAHYD
	i	ļ !		ļ		QQMGQKVQLPAETLQELLDLHRVSEREATEV
1	1	i I	ĺ	1 .		YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
						QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
1	1		l	Ī		PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
}						QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
1						VKAESAQASAKMVEEMQIKYQQMMEEKEKS
ļ	}			ļ		YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
		}			*	QEQARVLKERCQGESTQLQNEIQKLQKTLKK
12.10						KTKRYMSHKLKI
1 13/12	2608	T .	10404	1.5	1 802	TOT DADI COM CDI OL COCADI I TUMORTAC
1348	2698	Ā	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
1348	2698	Ā	10404	5	892	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS
1348	2698	Ā	10404	5	892	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
1348	2698	Ā	10404	5	892	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
1348	2698	Ā	10404	5	892	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM
1348	2698	Ā	10404	5	892	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1348		Ā	10404	5	892	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
1348	2698	Ā	10404	5	892	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1348	2698	A	10404	5	892	VAGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
1349	2698	A	10404	59	1184	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL
						VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
						VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ
						VAGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE
						VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV  LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
1349	2699	A	10409	59	1184	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM
						VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
1349	2699	A	10409	59	1184	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
1349	2699	A	10409	59	1184	VÄGGAPRRTTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKAQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
1349	2699	A	10409	59	1184	VÄGGAPRRRTPVTMWRLLARASAPLLRVPLS DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL RFIERAPLVPKVREPKNLSDIRGPSTEATEFT EGNFAILALGGGYLHWGHFEMMRLTINRSM DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE RERNNQNPWTFERIATANMLGIRKVLSPYDL THKGKYWGKFYMPKRV LRRNCSALGGLFQTIISDMKGSYPVWEDFINK AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLRQFSS ALIDCLINPLQEQMEEWKKVANQLDKDHAK EYKKARQEIKKKSSDTLKLQKKAKKGRGDIQ PQLDSALQDVNDKYLLLEETEKQAVRKALIE ERGRFCTFISMLRPVIEEEISMLGEITHLQTISE DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE PDPNGGGPTTASGPPAAAEEAQRPRSM AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW

## WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
  - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
  - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
  - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.



Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20